B16

PSYCHOLOGICAL TESTS APPLIED TO FACTORY WORKERS

BY
EMILY THORP BURR, A. M.

Reprinted from
ARCHIVES OF PSYCHOLOGY
Edited by R. S. WOODWORTH
No. 55

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, in the Faculty of Philosophy,
Columbia University

NEW YORK
May 1922

STANFORD UNIVERSITY LIBRARIES
WITHDRAWN
781
T72B
DECLASSED
Dissertation

PSYCHOLOGICAL TESTS APPLIED TO FACTORY WORKERS

BY

EMILY THORP BURR, A. M.

Reprinted from

ARCHIVES OF PSYCHOLOGY

Edited by R. S. WOODWORTH

No. 55

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, in the Faculty of Philosophy, Columbia University

NEW YORK

May 1922

Digitized by the Internet Archive
in 2025

https://archive.org/details/psychologicaltes0000emil

TABLE OF CONTENTS

CHAPTER I.

A Survey of the Psychological Testing of Factory Workers. 5

CHAPTER II.

The Factory and it's specific Problems. 14

CHAPTER III.

The Tests. 19

CHAPTER IV.

A Discussion of the Reactions to the Tests. 22

CHAPTER V.

The Correlations. 61

CHAPTER VI.

Multiple Correlations. 73

CHAPTER VII.

Interpretation of the Data. 81

CHAPTER VIII.

Confirmation of the Tests:—Three Years Later. 84

Appendix. 87

Bibliography. 92

PSYCHOLOGICAL TESTS APPLIED TO FACTORY WORKERS

Chapter I.

A SURVEY OF THE PSYCHOLOGICAL TESTING OF INDUSTRIAL WORKERS

It is rather disconcerting that in the conduct of modern business the handling of the human element has received far less consideration than the development of machinery and the perfection of the materials used in industries. The history of industry shows a gradual increase in the specification of the material factors employed, but not until the last decade has the recognition of differences in the individual been considered.

In the industrial magazines of former years a great deal more space is devoted to the study of the materials and machinery utilized than to the analysis of the employee and the appreciation of his needs and value. When selecting a machine, managers of industrial organizations are extremely careful to obtain one that is absolutely adapted to the specific work for which it is required, but although there is more variation in men than in machines, in the great majority of cases men are still selected by the most haphazard methods. It should be obvious that even though all factories producing a similar type of goods were to adopt the same kind of mechanical devices their output would vary tremendously because of the differences in the quality of the human labor employed.

Until recently, neither the employer nor the employee gave much attention to the method of work. The employer interested himself first of all in the Output and the wages he was obliged to pay while the employee concerned himself but little with the matter of Output and a great deal with the matter of wages. The attitude of both toward the amount of energy that must be expended to bring forth results was similarly one of more or less indifference.

There is consolation, however, in the fact that among progressive employers there are those who are becoming aware that the management and selection of employees is one of the most essential and at the same time one of the most complicated details with which they have to grapple. Unfortunately, few have known how to go about the curtailing of the waste which invariably results from chance selection and the large labor turnover which inevitably follows.

In times of business stagnation, there has existed a surplus of applicants for industrial positions and the employer has been rather undisturbed regarding changes in his personnel. During seasons when the wheels of industry are moving at maximum speed, the undesirables have had to be employed as well as the trained and efficient. In normal times, however, when the selection of employees fitted to their task becomes a factor that influences the cost of manufacture, the employer lays greater stress upon scientific management and is more ready to lend an ear to the theory of proper selection of labor and to the methods and tests designed to determine its fitness or unfitness.

Psychologists who have observed the incapacity of workers who are suffering from some defect of the brain have tried vicariously to call the attention of employers to the futility of expecting certain results from individuals fundamentally unsuited for the tasks assigned to them, but although there is here and there an inclination to apply scientific methods in the selection of employees, as yet the field is being cultivated more in the pursuit of an ideal than in the interest of industry.

In order to judge of the fitness of an individual for a specific task, it is necessary to analyze not only the aptitude of the applicant for the work, but also the character of the work as well. Certain industries demand special dexterity or a degree of mental alertness quite at variance to that required by other concerns.

The importance of determining the aptitude of a man or woman for a particular job is thus being gradually recognized, but the industrial world has a long journey to travel ere it reaches the same efficiency in this direction as that which characterizes its use of machinery or of raw materials.

Prof. H. L. Hollingworth recommends that we make psychographs of the different jobs and determine the physical, psychological, intellectual, moral, social and temperamental character necessary for each. (1). This is a task to be undertaken in every trade and in every branch of the different industries so that the mental factors involved in each type of work may be analyzed and the varying degrees of ability which each demands be decided upon. An immense field for investigation of this sort lies open to the Experimental Psychologist and offers an opportunity for an almost limitless amount of research.

Among the first to recognize the importance of scientific research along lines that would aid in the industrial situation were Dr. J. Mc. K. Cattell (2); Prof. H. Munsterberg (3); Prof. H. L. Hollingworth (4); Prof. E. L. Thorndike

(5); Prof. G. M. Whipple (6); to whose work reference is made in this paper.

In 1910, a Vocational Bureau was organized in Boston for the purpose of studying the human factor in industrial occupations. This seems to have been the first definite effort to analyze trade conditions from the standpoint of both employer and employee. It was found that only incidental attention was given to the subject of personnel organizations. In general, employees were treated only as economic units, and because of the lack of co-operation and understanding between employer and employee the labor turnover was enormous. The Bureau found that only a few firms scattered over the country had separate departments for hiring, training and promoting the welfare of their employees and that no scientific attempt to place applicants in positions suited to their individual capacity was being made by any of the firms then investigated.

A year after the inception of the Bureau, fifty men who were in charge of the Help Departments in fifty large establishments were asked to meet and discuss the problems of the employee. As a result of this conference, a concerted effort was made to apply the knowledge of the Twentieth Century towards the improvement of industrial practices. (7).

One of the chief factors which has brought about a more enlightened attitude toward the scientific examination and analytical study of the man or woman for the job, is the success attending the application of psychological tests to army recruits. During the war, the aid of psychologists was enlisted to examine and classify the intelligence of recruits so as to assign them to the duties for which they were best suited. The clever and ingenious tests then arranged proved a swift method of measuring the mentality of a man. Their application is some one million seven hundred thousand cases constituted a calculation of mental capacity on a larger scale than had ever been undertaken. (8).

A Committee on Classification of Personnel was appointed and it developed a number of Trade Tests which were successfully utilized in placing soldiers and sailors who came from every State in the Union. The tests were of three kinds: Standardized Questions, either oral or written, Picture Tests and Performance Examinations (9) which were arranged to determine a man's knowledge of the various trades.

The material for these tests was secured by a thorough investigation of each trade in order to discover suitable questions and comprehensive and fairly simple problems. When a series of questions was finally selected they were standardized by trying them upon groups of experts, journeymen, aprentices and novices. Questions which could not be answered, or work which could not be performed by novices

served to select the apprentices, while problems which could only be solved by the expert were intended to eliminate the unskilled. Those in charge of the Army Tests claim that their great value lies in their uniformity of method, in the brevity of time required to give them, and in the fact that they are so well standardized, that they can be given by persons without special training. There were two kinds of group examinations, one for men able to read and write English fairly well, called Alpha, and the other for persons not able to read and write English called Beta.

The interest awakened by these investigations of human intelligence has led to the discussion of the applicability of psychological tests in the industrial and business world. The question of the extent to which the present measuring devices of the army can be applied directly to industry depends, however, as Dr. Chapman so wisely observes, upon whether the conditions of an industry coincide with those of the army. (10). What the results of the Army Tests have proved is that it is no longer necessary to take a chance on a man's intelligence. Furthermore, they have demonstrated to the managers of business concerns a method whereby they may eliminate human waste and increase the efficiency of their employees. Therein lies the secret of their appeal to the business manager.

The testing that was done prior to that in the army did not arouse very much interest. Thus in 1915, when the American Tobacco Company engaged Prof. Scott to conduct some tests for that concern, there were only a few casual notices of the experiment. (11). The firm wished to have a scientific selection of salesmen made from a group of thirty-six applicants. It was arranged that six managers should interview the applicants and rate each by his own method. The results show the unreliability of haphazard methods of appraisal. The managers did not agree as to whether twenty-eight of the thirty-six applicants belonged in the upper or lower half of the group. One of the men was rated as Number One and as Number Thirty-two, another as Number Thirty and as Number Three, which shows the lack of uniformity in personal judgments.

John M. Bruce of the American Tobacco Company has applied certain standard tests that have proved useful in that industry. He claims that they determine "fitness, imagination, discernment, tenacity, address and personality." Each test is given a definite stated value, and the employee is rated according to his deviation from the average standing of the group. (12).

Thirty of the Efficiency Experts employed by the Cheney Silk Manufacturers were examined by Prof. Scott and the results of the tests correlated with the rating made by their

supervisors gave an index of plus .87. Cheney Bros. now use a series of psychological tests when interviewing applicants for high grade clerical and executive position, or when contemplating promotions. (13).

The firm of Joseph & Weiss rely in a large measure upon the results obtained from psychological tests in their selection of both executives and operatives. They especially emphasize the value of the Directions Tests and state that these tests check up accurately in nearly every case with the status of the men in the ranking made by their bosses. (14).

The Dallas Consolidated Railway Company (15) used a series of tests which was composed of questions and a device designed specifically to test the powers of observation of the workmen. It consists of a contrivance representing a trolley track and various obstructions to traffic. The test is based on Prof. Hugo Munsterberg's method for selecting Sea Captains and is supposed, in this case, to determine a man's ability to think and act with speed and good judgment in an emergency. They use also a set of cards which are to be sorted according to the letter which predominates.

These methods of investigation have proved so valuable in the selection of their employees that the length of service of all train-men has been increased seventy-seven percent in less than five years.

The Curtis Publishing Company also apply tests arranged by Prof. Munsterberg as supplementary evidence in determining the efficiency and accuracy of their stenographers and clerks. (16).

The Bell Telephone Company have used Prof. Munsterberg's tests to determine the degree of dexterity, memory, attention, accuracy and intelligence possessed by would-be employees, and report satisfactory results. (17).

In his excellent book entitled "Hiring the Worker" Roy Wilmarth Kelly discusses various applications of psychological tests. Mr. Kelly found that eighteen of the thirty firms to whom he sent questionnaires had separate departments for hiring employees and that all save two had been established about 1912. Each of these Employment Departments had analyzed the tasks demanded of the worker, had recorded findings, and were endeavoring to solve the problem of the labor turnover. He states that only three of the firms were using psychological tests. One firm was much interested in the results he had obtained from three tests. Another firm had used tests to avoid hiring those who were mentally unfit, and another firm had found tests useful in preventing the promotion of the mentally incapable. In all these cases the tests were employed not as vocational guides but to discover the limitations of the individuals. One firm, to be sure, was formulating some psychological aids to assist

in the selection of employees. They were using a General Intelligence Test, a Card-Sorting Test, a Word-Building Test, a Trabue Completion Test and a Cancellation Test. No data regarding results or methods was given. (18).

Some good work in the application of scientific methods to sales-management has been done by St. Elmo Lewis who lays particular stress upon the psychological elements involved. He criticizes the fact that too few managers analyze with precision both the job and the ability of the man who is to execute it, and urges the necessity for the two operations. (19).

Wm. F. Kemble in his recent book, "Choosing Employes by Test," sets forth very clearly that painstaking analysis is needed to differentiate between men who are engaged in even as simple matter as lifting weights, and shows how time may be saved by determining which men are adapted to handling specific degrees of weight. (20).

One of the most comprehensive studies of the use of psychological tests under working conditions in a representative industry is to be found in "Employment Psychology" by Henry C. Link. The first half of this book deals with the application of psychology to the selection of employes; the second half deals particularly with its application to the training of new employes, the keeping of records of their activities, and a discussion regarding methods of transfer and promotion. The entire trend of the work described is a scientific attempt to attain the applicant's view-point. He urges the use of tests which will discover the applicant's potential skill, his innate ability, and his general intelligence and the amount of his experience. He gives accurate directions for the presentation of a number of tests and the method of marking and recording results. (21)

In an article in "Industrial Management" for April 1920, Dr. Spaeth tells a story of a manager of a large plant who, as the result of the presentation of tests for general intelligence, discovered among his employes many with "unsuspected brains"! These fortunate "finds" he immediately advanced to the benefit of both factory and the individuals. Dr. Spaeth quotes the results of a few tests made on a large number of girls engaged in the inspection of shells for rifle cartridges. The "inspectors" who were recommended because of their excellent standing in the tests worked much more intelligently and for longer periods than did those girls who received marks below the average in tests. (22).

Myer Bloomfield has analyzed recently the routine duties of the employes in one of the largest and most progressive firms in Newark, New Jersey. Changes have been made in

the firm's organization as a result of his study of both the employes and the executives. He analyzed all the different tasks of the concern and arranged tests to fit each one. Dr. Bloomfield is now connected with the Vocational Bureau of Boston and is carrying on work there of a similar character in a larger field. (23).

During the year 1917, F. F. Taylor of the Metropolitan Life Insurance Company formulated a questionnaire which he presented to the employes. The results give an insight into the individual characteristics of the men, and proved advantageous in reorganizing the personnel. (24).

Mons. Jaques has excellent success in choosing typists and stenographers by means of pschological tests and reports high correlations with the subsequent standing of the girls in their actual output. (25).

In the Manhattan Trade School a series of tests is administered to see whether a girl has sufficient mental capacity to make her intensive manual training advisable. The Woodworth & Wells Directions Tests and the Easy Opposites are among the tests usually given. (26). Dr. K. Murdock has completed a Sewing Scale by which the work done by these girls may be graded.

Some of the findings of the Committee on Employment appointed by the National Association of Corporation Schools are of interest because of their reference to psychological testing. In their report for 1919 it is stated that "the number of instances in which tests have been successfully applied warrants the belief that their use can be extended to other industries." They announce also that the "purpose of psychological tests is conservative and their technique sane," and add that the success of any application of them depends wholly "upon the seriousness and thoroughness with which the testing is done." (27).

Any resume of the application of psychological tests to workers in industry would be incomplete if it did not include some account of the tests conducted by Dr. Helen T. Woolley in the city of Cincinnati. To quote from her own statement, she says, "The Bureau of Vocational Guidance has for its ultimate object the comprehensive study of the problem of child labor with a view to furnishing scientific demonstration of the effect of labor upon children who enter industry early." The children tested were a group of eight hundred boys and girls, over fourteen years of age who came to the Working Certificates Office during the year 1911-1912 to secure working papers. They were tested in serial order whenever they chanced to apply at the office. The scheme made provision that each of the children should be retested mentally and physically each of five succeeding years. (28).

In connection with the study made by Dr. Woolley, the Chamber of Commerce of Cincinnati, under the direction of Supt. H. C. Storm, has undertaken a survey of the general industrial situation of that city. He has organized the Public School System so that it may co-operate with the shops especially in reference to the printing industry which is an important one there. The work is still in an experimental stage, and the results have not been published.

Bernard Muscio in his "Lectures on Industrial Psychology", discusses the question of the desirability of the application of psychology to industry. He demonstrates that by adapting psychological methods to industry it has been found possible to obtain a given output from a smaller expenditure of human energy than has hitherto been found necessary, and he urges Psychology to turn to its laboratories again in order to better what has been done and devise practical ways for furthering its application to industrial work. (29).

It seems probable that much of the social unrest in the world is due to the misfits in the industrial world. The worker's lack of interest in the task assigned to him, and often almost total inaptitude for the work makes him a drifter. The concensus of opinion on industrial conditions is that it costs from fifty to two hundred dollars to train a person for a position. Each individual manager is well aware of the costly and discouraging features involved in the constant changes of his personnel. But many of these managers must still be convinced that, at least to a large degree, this bugbear of all industrial management, the labor turn-over, can be cut down by the exercise of greater care in engaging employees who are suited to their jobs.

The history of Vocational Psychology has been one of trial and error, and it needs must be conducted empirically for a long time to come. In its very essence it is made up of variable elements for some jobs change character almost from day to day, and the instability of the human element is readily conceded. In any review, then, that one makes of the subject of psychological testing one should note that the Vocational Tests are still for the most part in the experimental stage. There are no cut and dried methods that apply to all industries alike. Each one must solve its own problems. (30).

Tests have been standardized successfully in only about seven special lines of work other than "trades." The tests for the "trying out" of typists, stenographers, correspondents, labellers, clerical workers, inspectors, seamstresses and salesmen have all proved satisfactory and the norms established in each are reliable. (31).

The Trade Tests that were being standardized for use in the Army under the direction of The Committee on the Classification of Personnel have not yet been published. They were devised for some seventy-five or a hundred different types of trades and will doubtless be of greatest value when given to the public in the interest of vocational guidance. Dr. Chapman has described certain tests applicable in about thirty trades, and in his book entitled, "Trade Tests" gives accurate directions for their use. (10).

The study described in this paper was undertaken to demonstrate the value to be derived from the application of psychological tests in a hitherto unexplored field of industry. In addition to performing this function, the investigation is destined to give us information regarding the general intelligence of a certain group of women and girls who are earning a livelihood in industrial work. The mentality of delinquent women has been subjected to painstaking examination, (32), but the mental development of women factory hands, none of whom have failed to adjust themselves to social requirements, has received inadequate attention.

Chapter II.

THE FACTORY AND ITS SPECIFIC PROBLEMS.

Every time it is definitely proved that the reactions to any test, or to any group of tests, provide a means for the measurement of the capacity of candidates for a certain type of industry, a distinct service is rendered, not only to the employer, but to all potential employees.

The Experimenter secured the co-operation of E. Eisemann & Company, owners of a Feather and Fancy Ornament Factory situated at 63 West 38th Street, New York City, and received permission to submit their employees to a series of psychological tests.

Several days were spent in observing the work of the factory and studying the operations in the different departments. The Experimenter, although not actually serving at the bench, tried out each type of work herself so that the analysis might be as accurate as possible. Various members of the firm, as well as the heads of departments and the factory hands themselves, were interviewed and the Experimenter was able to profit by many of the suggestions they made.

The Manager said that he had relied largely, when engaging employees, upon the impression made upon him by the general appearance of the applicant and her replies to his questions, but he added that he would save time if he could have a more scientific method on which to base his decisions. He recognized the fact that his method did not permit him to discover natural aptitudes and he therefore formed his concepts of character and ability from observation alone. He would be glad, he said, to utilize a series of tests by which he might sort out applicants for employment, or use when promotions were to be made.

Before analyzing the characteristics which employees in this factory should possess, a careful study of each type of work had to be made.

The factory is divided into six distinct departments as follows:

FANCY DEPARTMENT
SELECTING DEPARTMENT
BENDING & FINISHING DEPARTMENT

BOA DEPARTMENT
DYEING DEPARTMENT
ERRAND DEPARTMENT

The work in the "Fancy Department" consists of sewing, glueing and assembling in endless combinations all sorts of feather materials in duplication of a presented model. This necessitates the exercise of judgment, discrimination, the ability to follow instructions and dexterity both with the needle and with the fingers. If a girl be imbued with imagination, initiative and creative ability, she is encouraged to originate models of her own design. In general, a higher order of mentality than that required in the other departments is demanded of the workers in this field, and consequently these girls receive the highest salaries.

The greatest number of girls is employed in the "Selecting Department" where the task consists of sorting the feathers according to their length and width. This requires the exercise of the powers of concentration and observation and demands a rapid visual estimate which must be accurate. A fair degree of discrimination is needed. This work is of a monotonous nature and involves the continual exercise of the same muscels of the body and of the same elements of the mind. Since the work is monotonous and not particularly interesting the labor turnover is greatest in this department. The work is divided into two sections, one which the raw feathers are sorted and the other in which the dyed, finished product is assembled.

The "Benders and Finishers" are engaged, as the designation implies, in the work of forcing the feathers into certain predetermined shapes. Manipulative skill, a great deal of patience and sufficient intelligence to follow instuctions are needed. Attention is the most important factor upon which demand is made.

In the "Boa Department" the work requires the same dexterity and persistency but under more trying circumstances since the bits of feather to be assembled are very small and accuracy in manipulation must not be sacrificed to speed. Girls of a nervous temperament are sorely tried by work of this character and rarely accomplish the same amount of work with the same amount of exactness as do fellow-laborers blessed with a more phlegmatic disposition.

The "Dyeing Department" employs more men than women and its secrets were not divulged beyond the fact that some knowledge of chemistry and a fair degree of intelligence are necessary requisites. Only two girls (167 & 117) from this department were tested.

In the "Errand Girl Division" the work consists mostly of fetching and carrying from one department to another.

The youngest and least experienced girls do this work, usually serving an apprenticeship as errand girls before being transferred to one of the other departments. It was, therefore, deemed of interest and importance to give them tests to determine for which type of factory work they were best suited. As Errand Girls per se, a fair memory and a spirit of willingness and co-operation as well as good comprehension are obviously the most essential characteristics demanded.

Racially, the workers might be divided, either by birth or descent, into four fairly distinct groups consisting of Austrians, Irish, Italians and Russians. The factory manager stated that he endeavored to keep these groups more or less distributed so that there would be slighter danger of labor union interference and the disturbance which would of necessity be caused when Jewish and other holidays occurred.

In reply to questions regarding their attendance at school, thirty-two of the subjects said that they had been graduated from the Public or Parochial Schools of New York City.

One girl had left school when in the Third Grade; three when in the Fourth Grade; two when in the Sixth Grade; seventeen when in the Seventh Grade; and eight when in Grade Eight. The remainder "could not remember" what grade they were in when they last attended school. Although the Experimenter was unable to verify these statements, there is no reason to doubt that they are in the main truthful.

TABLE ILLUSTRATING ABOVE SCHOOL HISTORY

Graduated of P. S. or Parochial Schools.....32
Left school when in grade indicated,

Third	1.
Fourth	3.
Fifth	2.
Sixth	7.
Seventh	17.
Eighth	8.
Did not recall	5.
Total No. of subjects	75.

When asked at what age they had last attended school, the replies indicate that the average age of those who left school before they were graduated was 14.5 years and the average age of those who did graduate was 13.73 years at the time of graduation. This suggests that the first group is composed of persons of a slower and less keen mentality than is to be found in the second group.

Many of the subjects regretted their lack of education yet said that they had been able to do their factory work fairly well and they seemed to like it in spite of its rather limited possibilities, that is they probably preferred work to school. They had all left school without having acquired any definite industrial or business training and it was by mere chance that they had drifted into this type of work. An introduction by relatives or friends already in the factory was the decisive factor. There had been no consideration of the suitability of the person for the job.

The subjects ranged in age from fourteen to fifty, but the majority were between sixteen and twenty-one years of age, as is indicated in this table:

Ages	14	15	16	17	18	19	20	21	22	23	24	25	26	36	40	50
No.	1	3	5	15	12	7	12	4	4	6	3	3	1	1	1	1

The spirit of the employed force was wholesome and pleasant. The Employer provided sick benefits, a service of free umbrellas and rubbers on rainy days, a small circulating library, and at more or less frequent intervals, social events which were arranged either by the firm directly or by the employees with the aid and support of the firm.

There were, therefore, no ulterior and difficult factors to be overcome before the testing was begun. No artifical or abnormal conditions that required adjustment seemed to exist.

While analyzing the jobs, the question of what tests to use was borne constantly in mind. There should be, it it seemed, an estimate of the level of intelligence from the standpoint of the individual and of the various groups engaged in the performance of the different types of work.

The Management emphasized especially the importance of ascertaining the capacity of a subject to comprehend and carry out instructions, so Directions Tests must be given to measure this particular ability.

In addition to these estimations, the perception, memory, judgment and reliability of the employee should be determined. She should be tested also as to her aptitude for making eye-hand co-ordinations. Her steadiness of control and concentration when a monotonous task is demanded of her should by all means be measured.

The degree of dexterity she possesses, her sense of color and proportion should in some way be demonstrated by the test-results.

The tests eventually chosen are described in the next chapter. Just what distinctive mental traits are denoted by an ability to perform this or that specific test cannot be definitely determined. There is too much over-lapping of mental qualifications to make this possible, but we can, in a measure, suggest some of the most important abilities demanded in the execution of the different tests.

Chapter III.

THE TESTS.

In order to establish test conditions as nearly standard as possible the co-operation of practically every one in the factory was sought and readily obtained. In addition, a spirit of keen interest in the tests and of friendliness toward the Experimenter was created which aided materially in the testing. The reasons for the investigation were explained to each one in turn. There was difficulty in gaining the confidence of but one of the subjects, an older woman, who felt that she was about to be submitted to some indignity that reflected upon her age and long connection with this business. She became interested, however, before she had progressed very far with the tests, finally confessed that she used to love school and wished she could have "gone higher." In the testing, therefore, little need be discounted on the score of nervousness or a lack of co-operation.

The seventy-five subjects who were examined by these tests and whose records are discussed later, spoke English readily and possessed a vocabulary, which if limited in scope, was yet used with perfect understanding and significance.

The Manager permitted the employees to leave their work for the tests, and since no recreation time was encroached upon, they were grateful for this break in the dull monotony of their daily tasks and reacted in a cheerful manner. In several cases, subjects expressed a positive delight in the tests which one of them described as "grand fun."

The tests were conducted in the Rest Room of the Factory and extended over a period of many weeks. The workers were examined singly, a proceeding that is probably unprecendented in the history of the testing of factory hands.

Before the actual testing was begun, the workers were made familiar with the method of examination by a series of preliminary tests the results of which were neither recorded nor used. This was done largely to eliminate the initial nervousness which fear of the "unknown" might have instilled in the hearts of some of the girls.

Each subject was given an identification number which appeared on all her test sheets. To distinguish between her identific number and the order given her in the firm's ranking, 100 was added to each identifying numeral. Thus

all numbers over 100 refer to the examiner's classification, while all numbers below 100 refer to the firm's classification.

A stop watch was used and the time required for the performance of each test was recorded.

The Experimenter found as did Dr. Link, (21), that unless the stop watch is concealed from the subject's view, she is inclined to be nervous and distracted by the idea that she is being timed. Only in certain tests were time-limits enforced. Limitations as to time would have defeated, in a measure, one of the secondary aims of this investigation, namely, to determine how the time records of these girls who are all engaged in factory work, compare with those of other types of subjects.

The directions were given orally because it was found that the girls follow spoken instructions much more successfully than written ones. The test sheets were kept face down until the subject was directed to begin.

Instead of devising many new tests, it seemed wise to employ tests in which norms could be used for comparative purposes. No detailed description of these tests is necessary inasmuch as they are widely known. The new tests which were arranged solely for use in this particular type of industry are discussed under the Sixth Group of Tests.

The mode of presentation and of scoring these Special Tests is in the Appendix.

The following is a list of the tests, presented in the order in which they were used:

I. General Intelligence — Binet-Simon
(Goddard Revision.)

II. Association Tests — Woodworth & Wells
A. Easy Opposites.
B. Mixed Relations No. 1.
C. Mixed Relations No. 2.
D. Substitution.

III. Cancellation Tests.
A. Digits Cancellation.
B. Letter Cancellation.

VI. Linguistic Tests.
A. Trabue Completion.
B. English-African Proverb Test-Ruger.

V. Directions Tests.
A. Easy Directions No. 1.
B. Easy Directions No. 2.
C. Hard Directions.

VI. Special Tests for this Industry.
- A. Judgment Test.
- B. Card Sorting Test.
- C. Sorting Test.
- D. Discrimination of Color Test.
- E. Motor Control Test.
- F. Feather Sorting.
- G. Weight and Form Test.

SPECIAL TESTS FOR THIS INDUSTRY.

These tests were arranged to call forth the different types of activity demanded in the several departments of the Factory.

Since the Selectors must judge feathers from the viewpoint of length, width and color as well as quality and suitability, Tests A., B., C., D. and F. should be especially desirable to aid in the discrimination of workers in this department.

The Boa and Stringer Department requires some ability to form judgments and a great deal of motor control, hence Tests A. and E. should be found useful indices of the best workers in this department.

For the Fancy Department Tests A. and D. seem to be indicated particularly.

The tests that require discrimination and quickness of perception may very profitably serve to demonstrate the degrees of proficiency attained by the various employees.

The Sorting Tests and the Discrimination of Weight and Form Tests should reveal the manual dexterity possessed by the girls.

The determination of the general intelligence of a girl, and her ability to comprehend and carry out instructions given her have been left to criteria established by the tests in which norms have already been ratified. The correlations derived from a comparison of the standing of the members of this group in the tests and their rating in the work of the factory is described in detail in Chapter Five.

The application of several tests has been employed in each case rather than the arbitrary presentation of a limited number of tests. Ability in one test compensates for deficiency in another, hence to present tests requiring different mental functions is the fairest way to form an estimate of the individual's capacity.

Chapter IV.

A DISCUSSION OF THE REACTIONS.

I. The Binet-Simon Test.

This test as arranged by Dr. Goddard (33) was presented in the prescribed manner to each of the seventy-five subjects.

The mental ages of the individual in the group range from 8.6 to 12. with the crude mode at 11.2. Twenty-nine fall below the mental age of 11 and only two attain a rating of 12. One employee passed but one test, (X-1), above the 8th year, and two others received a mental age of 9.4 and 9.6, respectively. The girls who had the lowest grades of mental development were found, with two exceptions, to be working on the boas, or bending feathers, or sorting them.

The distribution of mental-age on the Binet Scale may be observed by an inspection of this table in which the mental age is indicated and the percentage of subjects obtaining each age is given beneath it. The Median Age is 11: S. D. is 1.1358.

Mental Age	12	11.8	11.6	11.4	11.2	11	10.
% of Subjs.	3%	12%	11%	9%	18%	8%	11%

Mental Age	10.6	10.4	10.2	10.	9.6	9.4	8 8
% of Subjs.	11%	8%	3%	3%	1%	1%	1%

It is interesting to note how the results of this group, measured by the Binet-Simon Scale, compare with the records made by other groups.

Dr. Katherine B. Davis (34) in her Study of Prostitution in New York City writes that among 647 women at the Bedford Reformatory examined by the Binet tests, there were but two who measured 12 years mentally, twenty-six measured 11, and forty-four obtained a mental age of 10. The rest were below 10 mentally.

In this Bedford group, one-third of one percent had a perfect score compared to 2.66 per cent in our factory group; and while only four percent of the former group ranged in mental age from 11 to 11.8., inclusive, fifty-eight per cent. in

our group reached a mental age ranging from 11 to 11.8 Furthermore, whereas, 6.8 per cent reached a mental age of 10 to 11 in the Bedford Group, thirty-four percent received a like grading in the factory group. We may conclude, therefore, that this selected group of factory girls is of a higher type of mentality than the delinquents. Their opportunities along educational and social lines have been similar and one is led to deduce from a comparison of the results of these tests that the radical differences in the two groups are due in large measure to variations in native capacity.

The Binet-Simon tests conducted by Dr. Olga Bridgman and Dr. L. Morrow at the State School for Girls at Geneva, Illinois, show that but six of the girls are capable of passing the 12-year Test. Fourteen of the sixty tested were retarded from one to three years, and the remaining forty were retarded more than three years. The girls were all over sixteen years of age. That is eleven were retarded from four to five years and twenty-nine from six to thirteen years. (35).

In giving the Yerkes adaptation of the Binet-Simon Test to a group of mill-operatives, Dr. Yerkes found that errors were committed in all but three of twenty tests. In a Y. M. C. A. group, to whom the same tests were given, errors were made in all but five of the tests. Both of these groups consisted presumably of persons who had as good or better schooling than had the factory girls, yet errors were made in nearly every test. One can make no exact comparison between the Goddard revision and the Yerkes adaptation of the Binet, but this notation seems of interest inasmuch as basically the two tests are alike. (36).

In the "Report on Sensory and Mental Tests" made at the Idaho Industrial Training School, the author states that no attempt was made to apply any tests above the 12-year Scale of the Goddard Rivision of the Binet-Simon Test. All subjects who passed the 12-year Tests were counted normal. Eighty-three received a rating of 12., while fifty-five tested below twelve and were consequently called sub-normal. (37).

In his results obtained from a study of "The Measurements of Delinquents," T. H. Haines (38) states that among the one thousand individuals to whom he gave the Binet-Simon Test, five hundred sixty-three were rated below the 12th Year. In this connection, he speaks of the fact that some examiners, who use this scale, class those who measure below 12 as Feeble-minded, and he proceeds to disprove the method of classification that arbitrarily places the standard of demarcation at 12 years.

The rigid application of such a rule as this would place all save two of the Factory Group tested in this experiment in the sub-normal category, an obviously absurd classification.

Both Prof. E. L. Thorndike (39) and Dr. Stern (40) have shown the inaccuracy of the Binet rating above the age of 9 Years, and it must be conceded that to place the line of demarcation between the normal and the Feeble-minded at the 12-Year level on the Binet Scale is arbitrary and scientifically questionable. The amount of retardation that makes it impossible for an individual to manage her affairs varies with the individual and is to a large degree dependent upon her environment. In factory work, a girl of high intelligence is not necessarily the most desirable. One may note in this investigation that girls who have a mental age of 10 years, or thereabouts, are making good in this factory.

Let us next consider the individual tests of the Binet Scale, ranked in order of difficulty, beginning with the one in which most errors occurred.

This proved to be the repetition of sentences (XII.-3), in which 87 per cent. of the subjects failed. The dissected sentences came next with 62 per cent. incorrect; while 56 per cent. of the girls yielded to suggestion; 47 per cent. could not repeat seven digits (XI-1) in two out of three trials; 47 per cent lacked sufficient visual memory and analytical skill to reproduce the two designs (X-2); 44 per cent. were unable to discriminate between small differences in weight (IX-5); 29 could not repeat six digits two out of three trials (X-3); 12 per cent. failed to give definitions of "Charity," "Justice" and "Goodness" (XII-2); and 12 per cent. were unable to construct a sentence to meet the requirements of the Tenth Year (XI-5); 10 per cent. were unable to answer the Comprehension questions (XI-4); 9 per cent. gave the definitions of simple objects in functional terms only (IX-2); 6 per cent. did not give sixty free associations in three minutes (XI-3); 5 per cent. of the girls did not see the absurdities (XI-1); 2 per cent. could not repeat the names of the months without more than one error (XI-4); and one person (151) could not make change.

It has been stated that the subject finds the Binet test difficult in proportion to the time that has elapsed since his formal education ceased, but our results reveal nothing conclusive on this point. There are as many failures, comparatively, among those fresh from school as among the older girls and women who have been out of school many years. School Training to judge from an inspection of the table given below, has not influenced to any extent the rating on the Binet Scale of the members of this group. In this table, the employees are arranged on the basis of their grade on leaving school. The failures in the individual Binet Tests are indicated and the range in mental age of each group of those who left school when in a like grade are noted. It may be observed that the type of failures do not differ much from one group to

another. Thus, some public school graduates have a mental age of 10.2 while other subjects who left school from the 4th., 5th., or 6th. grade have received a mental age of 10.2; 11.2; 11.8 etc.

This peculiarity is in keeping with results obtained in the testing of other groups, notably the Army Group, where it was found that some of the highest intelligence records were made by men who stated that they had never completed the Elementary Grades.

It seems, therefore, that we are justified in concluding that the Binet-Simon Test is a measure of General Intelligence and native capacity which does not depend to any great extent upon school drill.

TABLE 3. THE BINET SIMON TEST RESULTS ANALYZED ON THE BASIS OF SCHOOL ATTENDANCE.

Group I. The thirty-two subjects in this group are graduates of Public or Parochial Schools. Their mental ages range from 10.2 to 12. The different tests in which these subjects failed to receive credit are noted below under the special test indicated:

IX	X				XI					XII					(Mental Age).
5	2	3	4	5	2	3	4	5		1	2	3	4	5	(Test Number).
4	8	5	2	1	6	1	1	11		12	1	18	15	7	(No. of Subjs.).

Group II. The seven subjects in this group left school in the Eighth Grade. Their errors are indicated as above. Their mental ages range from 10.6 to 11.8.

XI	X		XI			XII				(Mental Age)
5	2	3		1	2	3	4	5		(No. of Test).
3	6	1		2	2	6	6	2		(No. of Errors).

Group III. The nineteen subjects in this group all left school when in the Seventh Grade. Their errors are indicated in the same manner as are those of Grade I. The mental ages in this group range from 10.2 to 11.8.

IX	X	XI	XII	(Mental Age).
	2 3 4 5	3 5 1	2 3 4 5	(No. of Test).
	8 5 3 1	2 12 12	11 14 17 6	(No. of Errors).

Group IV. The twelve subjects who left school when in grades ranging from the Third through the Sixth Grade. Their mental ages range from 8.6 to 11.6., inclusive.

IX	X	XI	XII	(Mental Age).
5	2 3 4 5	2 3 4 5	1 2 3 4 5	(No. of Test).
7	6 3 3 3	2 5 1 7	6 1 9 5 2	(No. of Errors).

Group V. Five subjects could give no definite information regarding their school history. The mental ages of these subjects range from 9.6 to 11.8.

IX	X	XI	XII	(Mental Age).
5	2 3 4 5	1 2 3 4 5	1 2 3 4 5	(No. of Test).
2	3 2 1	1 2 1 3	1 2 2 4 2	(No. of Errors).

Since the only Auditory Memory Tests used were those included in the Binet Scale, a brief discussion of their results is of interest. Fifty-seven per cent. of the subjects have at least a seven digits memory span, and of this fifty-seven percent, two subjects passed all the Binet Tests; two obtained a mental age of 10.4; two of 10.6; three of 10.8 and three of 11 ; the remainder were about evenly divided between the mental ages of 11.2 through 11.8. Those who had but a six digits memory span comprised thirty-two per cent. These ersons obtained the following ratings on the Binet Scale: hree per cent. were rated 9.4; 10.4; and 11.6, respectively; and six per cent. of this six digits memory span group had a mental ge of 10.2 and the same per cent. of the six-digits group obained a mental age of 11.

The Five-Digits Memory Span Group consisted of eleven per cent. of the subjects, and they measured 8.6 (one only),

9.6 (one only), 10.4 (two only), 10.6 (two only), while the remainder were graded 10.8 and 11. mentally.

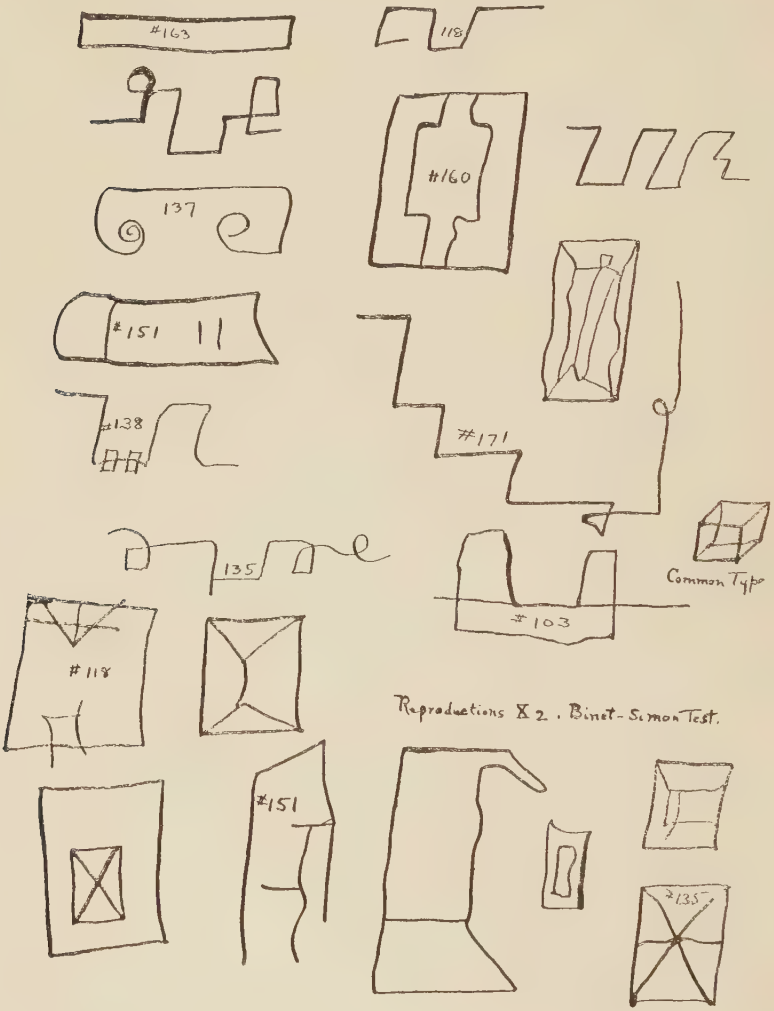
Thirty-eight per cent. of the Bedford Reformatory women have a Memory Span of not more than six digits; and thirty per cent. fail to recall even that number. (31).

In experiments with one hundred sixty-five college students, Gates (41) found that the average memory span for the group was 7.7 digits.

The result obtained from the auditory memory tests applied to this group of Factory Girls have been analyzed and given a Four Grade rating. Grade A denotes that the employe has at least a seven-digits memory span and has also been able to repeat after the Experimenter a sentence containing twenty-three syllables. Grade B signifies that the subject cannot repeat any one of the three sentences verbatim, but does possess the ability to repeat seven digits. Grade C shows that the subject's memory span is limited to six digits; and Grade D indicates that the subject has a five digits memory span only. This method of rating the auditory memory of a subject gives a basis for a fairly explicit type of group analysis that is useful in determining the concentration of employees in different departments of the factory.

Investigating still further the mental qualities requisite for workers in this factory, the reproduction of the two designs was found to have some "vocational value." It is an exercise that is quite on a par with the mental effort required from the workers in the Fancy Department in duplicating models submitted to them for reproduction. This is so true that those who failed utterly in this test were invariably found to be employed in other then the fancy Department.

The accompanying reproductions of these designs indicate how far astray some of the subjects went in visualizing the models presented.



It is obvious from an inspection of the following Table that employees who fail most decidedly in these tests are not disqualified for work in the Selecting, Boa and Bending Departments. It cannot be chance that no one employed in the Fancy Department has failed in the test.

Identification	Bank on Firm-Scale	Dept. of Factory
103	71	Boa-Stringer
107	35	Selecting
118	73	Boa
121	46	Selecting
128	69	Bender-Twister
135	45	Selecting
151	74	Selecting
160	67	Selecting
163	20	Selecting
164	18	Selecting
168	30	Selecting

WOODWORTH & WELLS ASSOCIATION TESTS.

I. Easy Opposites. (42).

This test was presented in the usual manner. The subject was asked if she understood what is meant by an "Opposite," and three illustrations were given her. She was then told to give the opposites of "Yes," "High" and "Wet." If the girl was unable to do this correctly, the directions were repeated and the words, "Big" and "Up" used as further illustrations.

In spite of these careful preliminary explanations, errors occurred which showed that the problem had not been properly digested. An analysis of the erroneous reactions reveals a failure to carry over an idea that in several previous cases had appeared to be perfectly understood, or else the inability to differentiate between words that sound alike yet have a different meaning and are spelled differently. There are still other instances in which an attempt at a definition has been made, and some reactions seem only explicable on the ground that associative processes were aroused by the stimulus word alone and the initial reaction was written without further thought.

"East" and "West" are reactions of this type, when given as opposites of "North." It is probable that had the subjects making these responses been catechized orally on this particular opposite each one would have known that South is the opposite of North. As an opposite of "Slow" one finds "Weak"; and as an opposite of "Weak" one finds "Thin," "Sick," "Brave" and "Well"; and evidently mistaking Weak for Week, some subjects gave "Day" or "Month." Attempting to define the terms, subjects gave "Tall" and "Big" as antonyms of "Large"; "Beyond" as an opposite of "Above,"

and "Delicate" as the opposite of "Good." Bua the mental proceses which prompted such reactions to "After" as "Now," "Seen," "Beyond," "Late," "Went," "For,"—are not easy to analyze. The opposite of "Love" resulted in many failures among which are "False Friend," "Mariage" and "Dont love you." "Come" brought forth such reactions as: "Went," "Going," "Stay," "Before," "Gone," "Don't be gone" and caused more errors than any of the other stimulus words.

No one neglected to give the correct opposites for "Rich" and "Dark," although they apparently are no less difficult seems probable that errors are due quite as much to a state of mental sluggishness as to a lack of linguistic knowledge. than some of the words that were frequently missed. It

The average reaction time for this group was 89.1 seconds; the average deviation in time was 25.92 seconds and the standard deviation in time was 34.75 seconds. The time range is from 29 seconds to 218."

The average score was 93.64 per cent. with an A.D. of 6.84 and S. D. of 30.45. The median score was 95 per cent. The range in score from 60 per cent. to 100 per cent. on the basis of 5 per cent. credit for each correct response. No half-credits allowed.

In the tests conducted by Dr. Bronner upon a group of delinquent girls, she reports a range of time in the Easy Opposites Test of 35" to 145", while in her Evening Settlement Group, the time range was from 37" to 130". In the Servant Girl Group, there was a time range of 52" to 215". The delinquent girls had no perfect scores, though some of the subjects missed only one-half of one opposite. (45).

The reports made by the Servant Girl Group in the Opposite Test show, according to Dr. Bronner, scores in which there were only 5., 6., 11., 12. and 13 correct opposites. In our group, the poorest performance was credited 60 per cent. or twelve opposites right, and 33.3 per cent. had perfect scores. Furthermore, Dr. Bronner permitted half-credits which has not been done in our scoring. While the entire lists of words were not exactly the same, they were of equal difficulty, approximately.

The probable reason for the prolonged reaction time of the Factory girls is due to the fact that they wrote their own responses. The girls in Dr. Bronner's group did not write their own reactions, and in the Easy Opposites Test conducted by Dr. Weidensall (32) at Bedford Reformatory, the reactions were written for the poorer half of the subjects. These discrepancies in method, unfortunately, make it impossible to offer any exact comparisons as to the time-records of these various groups.

2—The Substitution Test.

This test was presented in order to determine the rapidity with which new associations are formed by repetition. The practical problem is to combine the greatest possible speed with the smallest number of errors and consider both factors proportionately.

The time score is the chief guide because the mistakes are negligible,—six as a maximum, and a perfect score in the majority of cases. Yet our aim must be to find a formula which gives full value to both time and error and makes the material directly comparable in the form of one numerical value instead of two. In scoring, the method of penalties described by Prof. Munsterberg was employed. (46). By this method, the approximate range of both the quantitative and qualitative performance was found and then the units on the quantitative range were equated with the units on the qualitative range. The method equates time and speed, and by evaluating the results of the tests in this way a general conclusion as to the employees' efficiency may be made.

The average time score of the factory girls in this test was 176.5 seconds with an A. D. of 23.36., and a S. D. of 37.4". The time range was from 95" to 348.5".

Since the errors in this group ranged from 1 to 6, this makes the penalty for an error, 40.8" according to the method described above, and the employees were arranged in rank order of merit in the test on the basis of this penalization for errors.

When this test was presented by Woodworth and Wells to eleven adults the average time required for the test was 144.7 seconds, with a Standard Deviation of 12.5" and a range in time of 111 seconds to 176.5 seconds. (43).

The average time for the first half of the test in the Woodworth and Wells Group was 79.66 seconds in contrast to 92.4" required by the Factory girls to do a similar portion. The second half of the test was performed by the Woodworth and Wells Group in 65.1", while the factory girls required on an average 84.2" to do the latter half of the test.

Dr. Weidensall (31) reports a greater time range than that of the factory group. The Bedford women had a time range of 181.16 seconds to 618.5 seconds. The average reaction time for the first half of the test was 176.83 seconds and for the second half was 118.23 seconds. One is, therefore, justified in concluding that new associations are formed more slowly by the delinquents than by this particular group of factory workers, who, in turn, required a longer time and revealed a wider time range than did the adult group examined by Prof. Woodworth and Dr. Wells.

B. Mixed Relationship Test No. 1

Before presenting this test, the Experimenter assured herself that the subject understood fully what she was to do. Three examples were given her, outside of course, of those submitted in the examination; thus, Knee is to leg as Elbow is to? Shoes are to feet as Hats are to? Chew with one's teeth and Sew with one's? In this particular, a deviation was made from the regulation procedure. It will be recalled also that in this and in all other tests, directions were given orally. During the trial testing, the subjects appeared unaccustomed to written instructions, and asserted that they never made written lists of commissions to be remembered, and rarely, if ever, received other than spoken orders. This change in method seemed, therefore, to be wholly justifiable.

Inasmuch as this test requires more linguistic knowledge than the giving of Easy Opposites, the scores obtained are correspondingly lower.

The average score was 61.11 per cent.; the average deviation was 37.62. There were three perfect scores. The median score was 65 per cent. Scores ranged from 10 per cent. to 100 per cent. The Standard Deviation was 44.82.

The average time required for the test was 196.7"; the A. D. was 53.45; the S. D. was 70.73 seconds. The time ranged from 93 seconds to 400.4 seconds.

Woodworth and Wells give an average reaction time for a single response of 3.14 seconds, while the time required for a single test in the factory group was approximately 6.83 seconds. The subjects in the former group were college students habituated to intellectual exercises, and, therefore, scarcely comparable to the factory girls in a test of this kind. (44).

The errors that occurred in this test, analyzed, reveal an inability to adjust ideas in a new and unwonted manner. The failure to "carry over" a conception from one situation to another was, as in the previous test, the most frequent type of mistake. The repetition of one or another of the three words given instead of a reaction to the third word only often occurred. In a few instances, perhaps simply to avoid leaving a blank space, words that had appeared previously were written. Some of the reactions can be accounted for in no other fashion.

A few illustrations of the typical errors are given below; the words supplied are bracketted.

Eagle—Bird—Shark ..... (Bird, Snake).
Eat—Bread—Drink ..... (Starve, Glass, Eat).
Fruit—Orange—Vegetable ..... (Eat, Meat, Orange).

Sit—Chair—Sleep .....(Well, Walk, Lay, Lie, Rest, Night, Awake, Asleep).
Double—Two—Triple .....(Double, Single, Twice, Fish, Babies).
England—London—France .....(Germany, Italy, Spain, French, State, City, Country).
Chew—Teeth—Smell .....(Chew, Mouth, Gum, Odor, Flowers, Taste, Perfume, Gas, Good).
Pen—Write—Knife .....(Knife, Fork, Tool, Write, Sharp, Sight).
Water—Wet—Fire .....(Water, Out, Red, Sticks, Shrink).
He—Him—She .....(He, Him, Then, It, Woman).
Boat—Water—Train— ....Car, Engine, Elevator, Machine, Runs).
Crawl—Snake—Swim .....(Snake, Water, River, Bath, Swam, Walk, Float, Flew, Quick).
Horse—Colt—Cow .....(Colt, Kein, Dumb Animal, Bull, Pig, Dog, Milk, Chew).
Nose—Face—Toe .....(Finger, Face, Head).
Bad—Worse—Good .....(Bad, Best, Pure, Well, Excellent).
Hungry—Food—Thirsty .....(Hungry, Starved, Dry).
Hat—Head—Glove .....(Glove, Head, Shoe, Wear).
Ship—Captain—Army ....(Navy, Leader, Battle, Regiment, Men, Land).
Man—Woman—Boy .....(Woman, Male, Child, Men).

C Mixed Relations No. 2

This test appeared as No. 1. in the Monograph (14), but is here designated as No. 2 inasmuch as it was given after the other Mixed Relations Test numbered 2 in the Monograph. The results are similar both in time and score.

The average reaction time was 192.4 seconds; the A. D. was 56.25 seconds; the S. D. was 72.38 seconds. There was a time range of 66 seconds to 330 seconds.

The average score was 63.76 per cent.; the A. D. in score was 18.36; the S. D. was 29.46; the median was 65 per cent. The crude mode, or the score that was obtained by the majority was 85 per cent. The scores ranged from 100 per cent. to 15 per cent.

Instead of giving a list of the errors as was done in the previous tests, the different types of erroneous reactions have been analyzed; thus,

1. Repetition of one of the other of the given terms, as, Ear—Hear—Eye—Hear; or Hour—Minute—"Hour."

2. The use of a term that is descriptive of the third word as, H a m m e r—Tool—Dictionary—(Words—E ncyclopedia—Meaning Describes).

3. The introduction of a word that bears a certain relationship to the third word but is not the relation demanded, as, Wash—Face—Sweep—(Clean Broom); or (Sky—Blue Grass —Grows—Ground—In The Park).

4. The insertion of a word which suggests the completion of the thought as in this case.—Buy—Sell Come—(Here—Listen In— Home—Away).

5. There was noticeable also a tendency to insert a rhymed-associate to the third word, as, Straw—Hat—Leather—"Feather": and Hammer—Tool—Dictionary—Tarry: and Little—Less—Much—"Such": and Once—One—T w i c e—"Nice"; Oyster—Shell—Banana—"Anna": and Sky—Blue—Grass—"Glass": and Buy—Sell—Come—"Some."

It is difficult to determine by what process of reasoning this type of reaction was arrived at, inasmuch as the only correct verbal relation that rhymes is the first one i. e. Eye —See—Ear—"Hear" and evidently this first impression "held over." There was, however, no consistency shown by the subjects making these reactions, none of them attempted to find rhymes for all the words supplied.

6. Other failures were due to a lack of linguistic knowledge or an endeavor to define a stimulus word instead of fulfilling the conditions of the test, as., for instance, Little—Less—Much —"Let—Not Enough."

Many of these reactions reveal a groping for the right word, but a failure to grasp the relationship between the first two words clearly enough to reproduce it in an analogous case.

In this type of test the speed of the reaction depends on maintaining the proper adjustment throughout the series. Extraneous ideas must not be permitted to enter. Confusion, therefore, as shown in the failure to "carry over" the crux of the problem, indicates a lack of control and suggests that the subject who cannot form logical connections between ideas is inefficient.

Cancellation of Digits—The figure 2.

In this test the subject is required to cross off the figure "2" in a prepared form which contains this as well as many other digits. The methodology was the customary simple explanation of the problem, and the admonition to "begin at the first line and cross off as rapidly and carefully as possible all the figure "2s." The stop-watch was started when the subject turned over the sheet and stopped when she finished the last line. Some subjects traversed alternate lines from right to left. Two or three attempted to cross off

the digits along the vertical rather than the horizontal lines. This tended to increase instead of lessen the time, involved a waste of energy, and because of a diffusion of attention resulted in many omission of digits. The fact that one's eyes move from left to right in reading makes this the more natural movement in performing this test.

Five subjects complained of eye-strain as the result of the close application required by the task. The time consumed by these subjects did not vary materially from that of the average performance.

The average time, regardless of errors, was 201.5 seconds. The time ranged from 330 seconds to 120 seconds. Eight persons had a perfect score. One subject omitted seventeen twos. The average number of errors was 4.5.

The same system of penalties for errors was employed as that described in the Substitution Test. On the basis of the results of this scoring method, the subjects were arranged in rank order so that comparisons with other test records could be made.

The fact that the more intelligent of the subjects failed to make a high score in either time or accuracy may be due to the rather uninteresting material of the test. The subjects of a lower grade of mentality seized the test with avidity as if saying, as one girl actually did: "Here is something I can do easy." The task would seem to throw some light on whether or not a person is willing to perform routine work .

B. Cancellation of the Letter A.

There were more errors of omission in this than in the previous test. Two persons failed to cross off 32 of the A-s, one omitted 49! The number of errors ranged from one, then, to 49; while thirteen subjects had perfect scores. The average number of errors was nine. The median score on a percentage basis deducting one for every error irrespective of the time required was 90 per cent.; 100 per cent. perfect. The A. D. was 6.95.

The average time scored was 132.5 seconds. There was a range in time of 295 seconds: the best time scored was 68 seconds, the longest time required was 363 seconds. The A. D. in time was 35.19 seconds.

According to the Columbia University records, the A Cancellation Test requires, on an average, 95 seconds. (47).

The test has been given to various groups. Chambers reports the average time for the test as 75 seconds (48); Wallin as 100 seconds(49); while Doll (50), Burt (51), and

Wyatt (52) report the average time for the test as approximately 100 seconds. It is not surprising that the average time for the factory group is so much longer, because, in general, they react more slowly than do the normal members of group accustomed to doing a great deal of writing. The motor co-operations necessary showed among the older women disuse in a marked degree. For them, the task of writing their responses involves an actually laborious application of unaccustomed motor activity.

In reporting the results of this test in his experiments in testing twins (53), Prof. Thorndike did not count errors though they occurred in about one-third of the cases. The time element, he says, is the important factor in the test and accuracy may be made subservient to it.

In scoring the test, for the purposes of this investigation, however, it seemed expedient to use the method of penalties employed in the previous test. The subjects were given a rank order of merit rating on the basis of the speed and accuracy demonstrated by each individual.

Dr. Weidensall contrasts the performance of her group in this test with that of the Working Girls that Dr. Woolley examined in Cincinnati, and found the Bedford Group are slower in this simple task of checking letters than the working girls were. (31). Only a small per cent. of the Reformatory women were as able as the more efficient type of maid in cancelling letters with accuracy and speed. The scores made by the Factory Girls compared to those of the Bedford Group show marked differences in rate, the factory girls performing the task in a shorter time.

The Bedford Group have a time range of 610.4 seconds to 84.1 seconds which is 4 per cent. longer than the longest time record of the Factory Girls. The best record of the Bedford Group is 1.9 per cent. slower than the best record made by the girls at the Factory.

WOODWORTH-WELLS—DIRECTION TESTS.

The person who is able to carry out directions promptly and accurately is of great value in any business or industrial concern. This firm wished particularly to have a scientific estimate made of each employee in regard to this ability. The three tests arranged by Prof. R. S. Woodworth and F.L. Wells were therefore given. The two Easy Directions Tests were followed by the Hard Directions Test. (54).

The reading time of these tests varied greatly with the subjects, and although as the authors state, "The directions were made as concise as possible in order that the time re-

quired for the mere reading might not be a determining factor," yet in this group there were to be noted individual differences of wide range in the time consumed in reading and grasping the significance of the text. Let us consider each test in turn.

Easy Directions No. 1.

This set of twenty questions was performed correctly by forty per cent. of the subjects. The directions are so simple that there were few mistakes; a failure to react at all was more frequent than a wrong response. Thirty-eight per cent. of the errors were those of omission due either to carelessness or to a lack of comprehension of the problem involved.

Each direction is numbered from 1 to 20. The errors having been tabulated, the questions are arranged in order of their difficulty beginning with that question in which the most errors occurred, which happens, in this case to be the last, No. 20.

Number 20—Cross off the last word in this sentence.

Nineteen per cent. crossed off the last letter in the word Sentence, instead of following directions and crossing off the entire word "Sentence." One omission and a line drawn through the first word rather than the last were the only other errors in this test.

Number 6—Write the sum of these numbers 3—4.

This proved a stumbling block to eleven per cent. of the subjects who wrote one or the other of the following numbers: twelve, one-and-a-quarter, eight, nine, sixty, seventy-five or three-fourths. Nine per cent. omitted the problem.

Number 4. Write —|— over the longer word: It rained yesterday.

Ninety-one per cent. had a perfect score. Six per cent. omitted the question and three per cent. wrote one or the other of these words, "plus," or "over" or yesterday."

Number 3. How many feet makes a yard?

Eighty-four per cent. reacted correctly. Eleven per cent. stated that there are twelve feet in a yard, one said "36" feet and the question was omitted by three per cent.

Number 13. Make two dots between these lines: (lines drawn here).

This was done in 92 per cent. of the cases. Four per cent. made but one dot, instead of two between the lines; while three per cent. put two dots outside of the lines and one per cent. did not do anything at all.

Number 12. Write g on the egg-shaped figure (three designs given.)

Ninety-four per cent. did this; three per cent. wrote "g" on the wrong figure, and the remaining 3 per cent. did nothing.

There were no errors in the other directions. The average score was 89.14 per cent.; the Median score was 90 per cent.; the A. D. was 3.84 and the S. D. was 7.93.

The average reaction time was 92.4 seconds; there was a time range of 55 seconds to 296.2 seconds. The A. D. was 27.6; the S. D. was 35.34.

The average reaction time required by the Woodworth-Wells-group of educated adults was 72 seconds, with limiting records of 46 and 114 seconds.

Easy Directions Test No. 2.
(Beginning "Cross out the smallest dot").

This test is of practically the same degree of difficulty as the preceding test.

As in the other test, errors of omission were more frequent than any other sort.

The directions are numbered from 1 to 20 inclusive. Those in which errors occurred are noted below in the order of their difficulty, beginning with that one in which the most errors are found.

14. Write O after the largest number: 3, 86, 12.

Thirty per cent. wrote O above 86, ten per cent. wrote it below. Two per cent. wrote O after the 3, and two per cent. crossed off the number 86, while two per cent. failed to react at all.

15. Mark the name of a large city: New York, painter.

Instead of doing this the very easiest and most obvious thing, twelve per cent. wrote one or the other of the following words: New York City, Paris Fashions, Pittsburg, England, Chicago.

7. Write any word of three letters.

Four per cent. wrote words of four letters such as White, Word, Good and Four. Others wrote correctly a variety of words of three letters,—the words chosen were: eat, two mar, eye, the, how, are, got, his, dot, any and cat. The last word was the most popular.

There were one or two errors in the remaining directions, but they are not significant nor indicative of the intelligence of the group.

The average score obtained by this group was 85.45 per cent. The median score was 83.72 per cent.; the crude mode was 100 per cent. The A. D. was 6.84 and the S. D. was 12.96.

The average time for the group was 161.9 seconds; the A. D. was 44.85 seconds, the S. D. was 56,78 seconds. The time range was from 44" to 235".

The average time for the Bedford Group was 151.9 seconds, with limiting records of 55 to 397.4 seconds. (31).

When the extreme simplicity of both this and the previous Directions Test is considered, the difficulty which these girls experience in handling written work is apparent. In observing the reactions of the various members of the group, the reading over of the directions seemed to consume the major part of the time. Dr. Weidensall states that the best quarter of her subjects at Bedford read over the text in 40", the median subject in 50", and the poorest quarter in 90". The factory subjects showed a tendency to read and reread the instructions, and although admonished to "work as rapidly as possible," this inability to grasp quickly the meaning of the printed words seemed to be the cause for the retarded and prolonged reaction times.

3. Hard Directions Test.

This test measures the ability to read complex sentences and understand them and appreciate, when alternate clauses are given, that a choice is to be made. It is more important that the idea of making a choice be comprehended than that the correct response be written. In scoring a system of allowing half credits was arranged so that the significance of this feature of the test might be evaluated properly.

An analysis of the errors made in the test follows. The "directions" have been numbered and are arranged in the order of their difficulty, those in which the most mistakes occurred placed in gradation from a direction in which 51 per cent. of the girls failed to one with only 8 per cent. failures.

10. "Write NO if 2 times 5 equals 10," was an injunction disobeyed by 51 per cent. of the subjects all of whom wrote "yes.."

8. "Give the wrong answer to this question, How many days are there in the week?" elicited a response of "7" from

48 per cent. There were five omissions and the remainder of the subjects wrote either a 6, 9, 5, 10, 8 or 2. The digits 6 and 9 were the favorites. It is to be noted that no one wrote 1, 3 or 4.

20. "Write the first letter of your first name and the last letter of your last name," proved extremely confusing. Eleven per cent. failed because they wrote either their whole name or else omitted writing anything at all. Thirty-five per cent. wrote the initial letters of both their first and last name and could, therefore, receive but half-credit.

14. and 15. These two directions, the one dependent upon the other can best be discussed together: "Notice these two numbers: 3; 5. If iron is heavier than water, write the larger number here.....,"but if iron is lighter write the smaller number here.....

A subject, who recognizing the force of the alternate questions, as, for instance, a subject who indicates that water is heavier than iron because steamers are made of iron and float, should receive more credit than one who makes no choice. Some subjects placed numbers in both of the blank spaces, some put one or the other number in either space. Since this type of response showed a total lack of comprehension of the idea of choice both questions were called failures. One-third of the girls were unable to grasp the idea of choice and failed in both, and three per cent. omitted to fill in either space.

17. and 18. "Give the correct answer to this question: Does water run uphill?....and repeat answer here....." Twenty-five per cent. either omitted any reply or else wrote the word "down-hill" in the first blank only, and did not repeat the aswer.

19. "Do nothing here (5 plus 7 equals....), unless you skipped the preceding question." This question came next in order of difficulty, with a twenty-one per cent. of failures. Five and seven were added, or else, ignoring the significance of the plus sign, the digits were one or the other of them copied in the space, or else the plus sign was interpreted as a sign of multiplification, and "35" appeared in the blank space.

11 and 12. "Now, if Tuesday comes after Monday, make two crosses here.....(11).; but if not, make circle here.... or else a square here...(12)." These alternate choices caused a good deal of confusion. Twenty per cent. failed to see that a choice must be made, and thoughtlessly, made the crosses and inserted a circle or a square or both in the blank spaces. Nine per cent. omitted the question entirely.

16. "Show by a cross when the nights are longer: in summer?.....or in winter?".....The popular error in this reply was to put a cross in both blanks, which was a com-

plete failure for 19 per cent. Twelve per cent. indicated that nights are longer in summer than in winter.

3. "Then, if Christmas comes in March, make a cross right here.....but if not, pass along to the next question, and tell where the sun rises.....(4)." Twelve per cent. made a cross in the first blank and also wrote in the second blank one or the other of these phrases: "in the sky," "in the morning," "in the north," "south," or "west"; six per cent. omitted to answer either part of direction, while 16 per cent., recognizing the idea of choice to be made, received half-credit.

7. "Write YES, no matter whether China is in Africa or not....." Sixteen per cent. wrote No, while 3 per cent. failed to write anything at all.

5 and 6. "If you believe that Edison discovered America, cross out what you just wrote (5), but if it was some one else, put in a number to complete this sentence: "A horse has ....feet." (6). No one crossed off what she had just written, but five per cent. did cross off the word "Edison," and in its stead wrote "Columbus." There were several interpolations of a number between the semicolon and the sentence, "A horse etc." The horse was given but two feet in 5 per cent. of the cases; however, there were only two omissions and hence 90 per cent. of the subjects reacted correctly to both 5 and 6.

9. Write any letter except g after this comma." This seems a very simple and straight-forward direction yet there were 6 per cent. of the subjects who proceeded to write g in spite of the injunction to do otherwise. Four per cent. wrote nothing.

2. "Put a comma after the longest of these three words: boy mother girl." There were types of failure: a comma was placed above the longest word, or after each one of the words, or after the word "girl," or nothing at all was done. The total number of failures amounted to 9 per cent.

1. "With your pencil make a dot over any one of these letters: F G H I J." This direction, proved the easiest to interpret, met with failure in 8 per cent. of the cases. Dots were placed over each one of the letters, while one girl dotted most of the given letters as they appeared in the text. There were no omissions.

The average score in this test was 77 per cent.; the A. D. was 9.41; the S. D. was 11.32; the median score was 80 per cent.; the scores ranged from 100 per cent. to 45 per cent.

The average time score was 262.12 seconds; the A. D. in time was 48. 58 seconds; the S. D. in time was 53.49";

there was a time range of 70 seconds to 719.5 seconds.

Woodworth and Wells quote 107 seconds as the average time for their group of college students. We can report but three subjects with as good a time score as this average made by the students, and Dr. Weidensall had only two subjects among the Subjects examined who require as short a time. Of the Reformatory women all but one were slower than 134 seconds which was the poorest time-score of the college students. (31). The average time for a group of sixty normal students at the Chicago Normal College is 118.8 seconds (31) which is slower than the Woodworth and Wells subjects, but a better time-record than was made by the majority of our subjects, since only the three mentioned above performed the test in as short a time. The index between time and accuracy derived from this test shows that the highest score was not obtained in the shortest time, nor on the other hand was the slowest person the most accurate. The relation between time and accuracy is discussed under a separate heading.

The Bedford Reformatory Group performed the test in 274.85 seconds, average time for the group, which is 12.63 seconds slower than the average time required by the Factory Girls. (31).

LINGUISTIC TESTS.

A. The Trabue Completion Test.

The Trabue Completion Test beginning "The sky..... blue." appeared to be an interesting form of amusement to most of the subjects, for they went smilingly through the test, making errors often but generally writing something in the blank spaces.

Since the English spoken by the majority of the Factory workers contains much incorrect phraseology, it seemed wise to overlook minor details in the results, and count as errors only impossible and meaningless interpolations. It was common to find the plural subject of a sentence completed by a verb in the singular, and vice versa. A logical connection of ideas rather than grammatical rectitude seems a fairer criterion than to lay undue stress upon linguistic construction. This group made some interpolations that are unique, and as they are not noted by other investigators, Dr. Trabue was consulted in regard to their proper valuation. In general, however, the prescribed method was utilized. A system of half-credits was found necessary.

The sentences vary so much in point of difficulty that they must be discussed individually.

1. The sky.....blue.

No one failed to interpolate a word that made sense and "looks", "was" and "is" were used with about equal frequency.

2. We are going....school.

The only errors in this sentence were the insertion of "home" and "play" by 3 per cent. of the subjects.

6. The.....is barking at the cat.

Every one had this correct.

8. The stars and the.....will shine to-night.

Four per cent. of the subjects inserted the word "sun" 21 per cent. supplied the word "stripes", while "moon" was written by the remaining 75 per cent. of the subjects.

10. The.....sings a song every morning.

The only error here was the use of the word "will".

12. Good boys.....kind to their sisters.

Here one finds "be", "run errands", "and so are" inserted instead of the very obvious "are" which was given correctly in 95 per cent. of the cases.

14. When the.....grows older he.....be a man.

Four per cent. inserted a proper name in the first blank over looking the significance of the definite article "the," otherwise there were grammatical errors only in this and the two succeeding sentences.

16. The boy will.....his hand if.....plays with fire.

The word "wash" was used in 8 per cent. of the cases.

18. The best.....to sleep is at night.

This was correct in 78 per cent. of the reactions, the remaining 22 per cent. inserted the word "thing", or "cure", "his", "is", "to go", "place", "way", or "hour", showing an inability to comprehend the meaning of the sentence.

20. The little.....and his dog.....running a race.

The only errors in this sentence were those in which a plural subject was followed by a singular verb and vice versa.

22. Time.....often more valuable.....than money.

Fifty-two per cent. had these two insertions right; the rest of the subjects wrote one or the other of these words "Work," "Come," "Has," or "Are" in the first space and "in" or "to" in the second space quite as often as Than.

24. The.....rises in the morning and.....night.

The words "Boy", "Girl" or "Bird" were supplied by 40 per cent. of the subjects "rising" in the morning and "sleeping" at night. The majority, however, wrote "Sun" correctly, while 50 per cent. inserted "sets". The remaining 50 per cent. wrote "lowers" or "moon" in the second blank.

26. Boys who play.....mud get their hands.....

These blanks were filled in correctly by 50 per cent. of the girls. There were various insertions descriptive of "mud", such as "nasty," "street," "black," "much," "bad," "soiled"

and "yard". In the second blank "bad", "full" and "soiled" were chosen in about equal proportions.

28. The little....had....nothing to....; he is hungry.

But 6 per cent. failed to recognize the significance of the pronoun "he" and wrote "girl" or "creatures" in the first blank.

30. The boy who.....hard.....do well.

This sentence was confusing to 70 per cent. of the subjects who inserted "have", "is", "tries", "has",, "studies", "eats" or "thinks" in the first blank, and "always" "work" or "does" in the second.

32. One's.....do.....always express his thoughts.

Eleven per cent. wrote the word "who" in the first space and fifteen per cent. the word "mind". Seventy-one per cent. supplied meaningful words, but the remaining 18 per cent inserted either "ideas," "opinion," "self," "actions," "speech" or face," "Not" was written invariably in the second blank space.

34. It is a.....task to be kind to every beggar..... for money.

Words that made sense were inserted by 33 per cent., while 50 per cent. wrote words that could be given half-credit. The remainder left the spaces blank.

36. Worry.....never improved a situation but has .....made conditions.

Thirty-five per cent. of the subjects completed this sentence correctly; twenty-five per cent. received half-credit and the remaining thirty-seven per cent. failed to insert any words that could be accepted; ten per cent. made no attempt to supply any words.

The average score was 80.33 per cent; the Median Score was 84 per cent. The scores ranged from 100 per cent. to 34.8 per cent. The A. D. in score was 13.76 per cent. The S. D. was 24.62 per cent.

The average reaction time was 464.5 seconds; the median time rate was 340.5 seconds; the time scores ranged from 293 seconds to 787 seconds.

LINGUISTIC TESTS—ENGLISH-AFRICAN PROVERB TEST.

No published data of this test arranged by Dr. Ruger of Columbia University has appeared. Prof. Thorndike included the test in a set given at Teachers' College, but no detailed report of the results of the test are noted. (55).

The Experimenter beside presenting it to this group of Factory girls, gave the test to a group of persons employed at the Willard Parker Hospital. The reactions of these two

groups may well be compared.

The directions were read to the subjects, and each one was asked in turn if she understood what was to be done. If she did not comprehend the problem, she was asked to read aloud the first of the English Proverbs, and then find among the African Proverbs the one that seemed to mean most nearly the same thing. Stimulated by a suggestion as to a method of procedure, the subjects invariably did something with the test.

The Factory subjects were asked after they had read over the test if there were any words which were unfamiliar. Six girls announced that they did not know "enchantment," and "Leisure" and "Folly" were new words to three girls. It is possible that other persons failed also to grasp the meaning of all the words but were too shy to admit their ignorance. It is obvious from an inspection of the accompanying table that some of the confusions occurred through a lack of comprehension of the phraseology.

Results of the Test administered to the Factory Girls are as follows:

5 per cent. had a perfect score.
5 per cent. had eleven of the analogies right.
1 per cent. had ten of the analogies correct.
3.33 per cent. had nine of the analogies correct.
1 per cent. had eight of the analogies correct.
5 per cent. had seven of the analogies correct.
8.33 per cent. had six of the analogies correct.
31 per cent. had five of the analogies correct.
8 per cent. had four of the analogies correct.
21 per cent. had three of the analogies correct.
10 per cent. had two of the analogies correct.
15 per cent. had one of the analogies correct.
5 per cent. had none of the analogies correct.

Each pair may now be discussed in the order in which the English Proverbs occur in the text.

A. Married in haste we repent at leisure.

11. Quick loving a woman means quick not loving a woman.

These two seemed by far the easiest proverbs to analyze and 87 per cent. paired them correctly; only 6 per cent. omitted A entirely, 3 per cent. paired 11 with "First catch your hare" (G), and 3 per cent. paired 11 with "Distance lends enchantment to the view" (K). There is a semblance of humor in the latter combination, unless one chooses to ascribe it to chance! Every one had heard the English version of the proverb and the expression offered no complication to most of the subjects. Except to the very stupid, the African interpretation of a hasty marriage was perfectly obvious.

B. Answer a fool according to his folly.

6. If the boy says he wants to tie the water with a string, ask him whether he means the water in the pot or the water in the lagoon.

This analogy was given by 34 per cent. of the subjects. Number 6 was paired with "Distance lends" etc. (K) by 9 per cent; and was coupled as well with A., C., G., H. and J., respectively, by 8 per cent. of the subjects; while 3 per cent. paired No. 6 with L. and 3 per cent. with M., and 4 per cent. coupled it with D.—"Out of the frying pan" etc. Practically one half of the subjects failed to make any combination with No. 6 and 8 per cent. made no attempt to dispose of B.

C. One swallow does not make a summer.

1. One tree does not make a forest.

Fifty-one per cent. had these two paired correctly. Twice No. 1 was combined with "Birds of a feather etc.", because as these subjects explained, both proverbs contain the word "Bird." Eighteen per cent. failed to pair C. with any other proverb, and twenty per cent. failed to find an analogy for Number 1.

D. Out of the frying pan into the fire.

8. He runs away from the sword and hides himself in the scabbard.

Only twenty-three per cent. had this correspondence correct. Fourteen per cent. combined E. and 8. Six per cent. thought J. and 8 analogous. Three per cent. paired 8 with I., and three per cent. paired it with M. and three per cent. with K. Twenty-one per cent. failed to make any combination for D., and twenty-nine per cent. omitted No. 8.

E. Robbing Peter to pay Paul.

13. No one should draw water from the spring in order to supply the river.

These two proverbs were arranged correctly by twenty-eight per cent. of the subjects. Ten per cent. paired M. with 13; and several paired it quite at random, apparently. J. was combined with 13 by 6 per cent. of the subjects, and D was combined with 13 by four per cent. Twenty-five per cent. omitted E. entirely, and thirty-one per cent. did nothing with 13.

F. Birds of a feather flock together.

9. A fool of Ika and an idiot of Lluka meet together to make friends.

Thirty-one per cent. had this analogy correct. Twenty-three per cent. paired 9 with B., "Because," said they, "both contain the word 'fool'; six per cent. paired 9 with I., arguing that for a fool and an idiot to meet is "adding insult to injury." Three per cent. joined 9 with A., "For," said these subjects, "persons marrying in haste" are a fool and an idiot meeting. Number 9 was combined with E. an equal number of times as "left over" propositions that could not be disposed of other-

wise. Fifteen per cent. failed to pair F. and seventy per cent. omitted Number 9.

G. First catch your hare.

2. I nearly killed the bird. No one can eat nearly in a stew.

This pair of proverbs proved more difficult to analyze than did any of the others. Number 2 was paired at least once with every other proverb, and with F. in seven per cent. of the cases. Both F. and E. contain the word "Bird" which was given as the justification for this arrangement. One person insisted that "No one can eat nearly in a stew" means that no one can eat when nearly in a fret or hurry—i. e. in a stew. Eighteen per cent. made the right combinations. Forty-two per cent. did not pair G. at all, and twenty- six per cent. left out Number 2.

H. Sour grapes.

7. Cocoanut is not good for bird to eat.

This pair appeared as difficult to analyze as the previous one inasmuch as the same percentage, (thirty-one), arranged this and the former pair correctly, although the combinations made were not as varied as in the C. and number 2 Proverbs. "Sour grapes" proved a fairly familiar saying, but "Cocoanut" suggested the idea of food, "Milk for babes" (L) to twenty-one per cent. of the subjects, and to six per cent. the word "Bird" made the combination of Number 7 with F. seem the logical one. Six per cent. of the subjects coupled Number 7 with E., G. and J. respectively. Eighteen per cent. made no disposal of H. and thirty-one per cent. none of Number 7.

1. Adding insult to injury.

10. The ground pig said: I do not feel so angry with the man who killed me as with the man who threw me to the ground afterwards.

Twenty-one per cent. found this analogy. Twelve per cent. paired Number 10 with J., and seven per cent. paired it with M., and seven per cent. with G. The last two proverbs refer to animals, the subjects who made these combinations remarked, so why not put them in pairs! Six per cent. coupled Number 10 with D., and three per cent. with B., and H. Fourteen per cent. failed to dispose of I. in any fashion and fifteen per cent. did nothing with Number 10.

J. Curses come home to roost.

3. Ashes fly in the face of him who throws them.

Only twelve per cent. of the subjects combined these two. Fifteen per cent. paired Number 5 with D., because the word "ashes" suggested "fire" and there was none other that seemed to "fit as well." Each one of the English Proverbs was paired with Number 5 at least once, and, as one girl expressed it, "I put down all I was sure were right, and the rest I just settled any way at all! This method of elimination was

quite the usual one. Thirty-seven per cent. left J. a blank and thirty-one per cent. omitted Number 5.

K. Distance lends enchantment to the view.

4. Distance firewood is good firewood.

Upon the supposition, perhaps, that proverbs containing the same word, or one of its derivatives, should be paired, twenty-five per cent, of the subjects combined these two immediately. Fourteen per cent. followed a similar line of reasoning, allied Number 4 with D., since "both have to do with fire." An equal number of persons thought Number 4 and F. were analogous, while in three per cent. of the cases Number 4 was paired with B., C., H., and M. Thirty-four per cent. omitted K entirely, and fifteen per cent. failed to pair Number 4.

L. Milk for babes.

12. If the stomach is not strong do not eat cockroaches.

These two proverbs are too abstruse, apparently, for this group since but ten per cent. grasped the analogous meanings. "Milk for babes" was paired with Number 3. in twenty-six per cent. of the cases, an association of ideas that can be understood. Twelve per cent. made no disposal of L., and thirty-six per cent. none of Number 12. Twelve per cent. combined Number 12 with H., while six per cent., recalling the pangs of indigestion, paired Number 12 with "Curses come home to roost" (J.), or as one girl expressed it, "If you eat what does not agree with you, you always have to pay!" Six per cent. reasoned that giving the "stomach cockroaches to eat" was analogous to "adding insult to injury" (I). There were other combinations that the introspections showed were studied, but the majority were chance associations only.

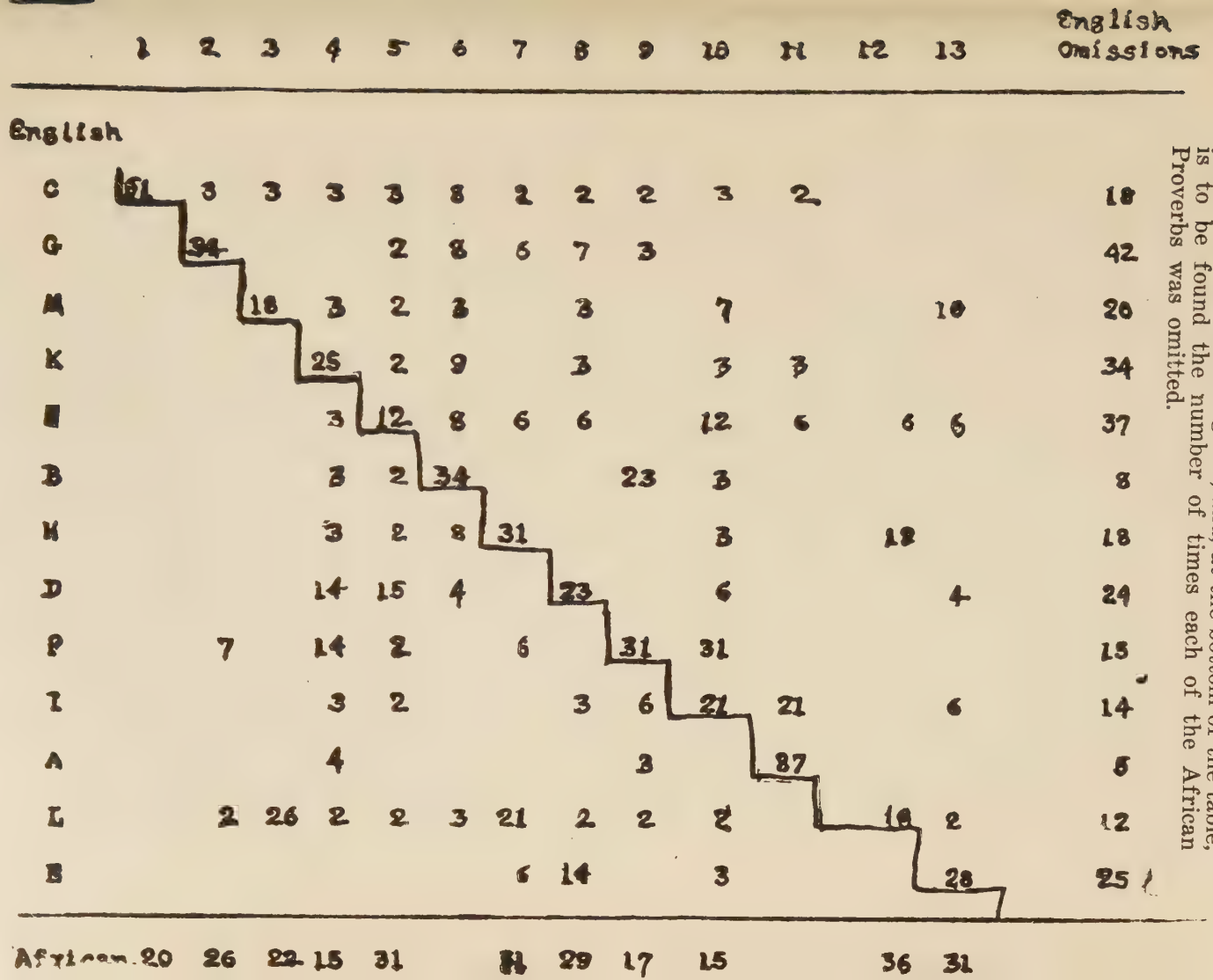
M. We can all endure the misfortunes of others.

3. Full-belly child says to hungry-belly child, "Keep good cheer."

Eighteen per cent. paired these two proverbs, but twenty-six per cent. combined Number 3. with L., and M. was coupled with every other proverb in the African Group, except Numbers 1 and 11, at least twice. Twenty per cent. failed to make any combination at all with M., and twenty-two per cent. failed to connect Number 3 with any other proverb.

Arranged in the order of difficulty, beginning with the proverb that was compared correctly most often, the English Proverbs appear thus:—A., C.,; B., H. and F.; I., K., and D.; G. and M.; J. and L.

The table given below reveals in a graphic manner the number of correct analogies and the omissions that were made, as well as the confusions and errors that occurred. These have been noted, but the tabulation gives a better idea of comparative difficulties in analysis than can be derived from a purely verbal elucidation. The zig-zag line through the center



designates the correct pairing of the proverbs. In the vertical column to the right, the number of times each of the English Proverb was omitted is given, and, at the bottom of the table, is to be found the number of times each of the African Proverbs was omitted.

It was noted that the first six persons who tried the test completed it in less than 13 minutes, the next six persons made neither addition nor correction after a period of thirteen minutes had elapsed so a time-limit of thirteen minutes was set. In the experiment conducted under the direction of Prof. E. L. Thorndike (55) an allowance of twenty minutes was made for this test in order that the measurement might be wholly of the quality of the work done irrespective of the time element. The Factory situation did not warrant the devotion of so much time to this one test. The Experimenter feels confident that the test-results would be practically the same had more time been allowed the factory girls. No time limit was fixed for the Hospital Group and their time ranged from 900 seconds to 185 seconds, with an average time record of 543.14 seconds; the A. D. was 153.39 seconds.

The time consumed in the execution of the test by the Factory Girls ranged from 780 seconds to 312 seconds. The average time for the test was 581.149 seconds; the A., D. was 153.39 seconds.

The test was scored on the basis of 100 per cent. perfect, or an allowance of 7.7 per cent. each correct answer.

The average score for the Factory Group was 29.5 per cent.; the A. D. was 6.45.

The average score for the Hospital Group was 72.94 per cent.; A. D. 18.93; the S. D. 26.54.

VI. DESCRIPTION OF THE SPECIAL TESTS.

1. Judgment Test.

The Foreman in the Selecting Department had observed considerable variation in the ability of the girls under his direction and wished to determine how they compared in efficiency as measured by a series of special tests, and whether any other girls possessed aptitude for the work. In will be recalled that girls in the Selecting Department are obliged to sort feathers according to their length and width, regardless of their quality. This is the most important work done in the department, since an error of a quarter of an inch in the estimation of the dimensions of a feather means quite a loss of money to the firm, hence to assist the inexperienced worker there is a foot-rule scale on each girl's table. She is expected to lay the feather on the scale if she cannot make an independent decision. The more exact and expeditious the sorter becomes the greater is her individual value to the firm, and the truly efficient sorter forgets that the scale exists. Now, although practice in this type of work is of the utmost importance, the aptitude for making quick decisions varies with the subject.

In order to determine the degree of variation in forming such judgments as are demanded in this department, the following test was arranged.

Twenty strips of fairly heavy card-board cut to approximate the sizes of the feathers examined during the ordinary routine of the day were prepared. (See measurements given in the Appendix.) These strips were numbered to correspond with the scale in actual use in the Shipping Department, so that they might be readily identified by the Examiner and checked up quickly.

Each strip of card-board was held at distance of eighteen inches from the subject and she was asked to estimate the length and width of the strip, writing her appraisement on the specially prepared sheet that had been supplied her for this purpose. The strips were exposed for ten seconds. Although rapidity of decision counted in favor of the person examined it did so only when quickness of answer coincided with exactness of answer. At no time was wrong answer rapidly arrived at counted as anything but a failure against the subject giving it. . . .

In scoring, 5 per cent. credit was given for each correct judgment, and for every error of half-an-inch in an estimation a deduction of one per cent. was made from the score.

Sixty-seven per cent. of the judgments of width were correct and fifty-two per cent. of the judgments of length.

The average error in judgment was 1.53 inches, with a range in error of 1.88 inches, or errors that ranged from .62 inches to 2.5 inches. . .

The average over-estimation of length was 1.83 inches, while the average under-estimation was 1.09 inches. The range of error in under-estimating the length of the strips was from .45 inches to 3.04 inches. The average error in over-estimating the length was from .63 inches to 2.53 inches.

In the matter of under-estimating widths there was an average error of .87 inches, and in over-estimating widths of .869 inches. The range of error in over-estimations of width was from .58 inches to 1.68 inches. All of which goes to show that there was a somewhat greater tendency to over-estimate than under-estimate the dimensions.

Ritter (56) reports in an experimental study that there is a definite inclination to over-estimate vertical distances and under-estimate horizontal distances. This may be due to the increased effort required in moving the eyes up and down instead of in the more accustomed left to right movements used in reading. The differences in judging the length and width of the feathers made by these girls are indicated in this table:

	Over-Est. of length	Under-Est. of length	Over-Est. of width	Under-Est. of width
Average	1.876 in.	1.09 in.	.869 in.	.777 in.
A. M. D.....	.611 in.	.534 in.	.188 in.	.428 in.

The time required for the entire test averaged five minutes and three seconds. The quickest person made her decisions in three minutes and three seconds and the slowest in seven minutes and twenty seconds.

The index between time and score was higher in this than in the purely intellectual tests, showing a correlation of .63 (positive), with a Probable Error of plus or minus .059.

It is possible as a result of the analysis of the records made in this and the two following tests to discriminate with much exactitude regarding the output of the individual worker.

As has been observed, the ability to form quick and accurate judgments in the matter of color, form and quality is a most valuable asset in a factory employee. The table showing the individual correlations of the girls and their rating by the Firm illustrates the practical application of this particular group of tests.

It was observed that those persons who hesitated over their decisions were the ones who made the greatest number of errors; but, on the other hand, those who passed judgments with scarcely a glance made more errors than the subject in the Median Time Group who are neither slow, nor yet too hasty.

There are a preponderating number of over-estimations of both length and width, showing what might be interpreted as a tendency toward exaggeration. The Manager, noting this tendency when inspecting the results, stated that he wished there were tests to reveal the presence of this trait, for among the "outside men" employed by the Firm unnecessary difficult arose because either they over-estimated the promptness with which orders could be executed, or else they exaggerated the quality and character of the goods.

II. Sorting Test.

A hundred strips of fairly heavy card-board were cut to correspond with twenty of the most frequently used feathers such as were described in the previous test. There were, therefore, five strips of each size. These strips were laid on the table in front of the subject in a heterogeneous mass just as the feathers are received from the stock-room. Each girl was directed to sort the strips according to size, treating them as though they were feathers and laying those of like dimensions in the same pile. They were told to work as quickly and carefully as possible.

This test requires the exercise of the same sort of judgment, concentration and motivation that is involved in the performance of such a piece of factory activity as this: suppose that an order for a hundred feather fans has been received, and several thousand feathers are sent down from the stock-room each sorter is given a share and instructed to lay feathers of the same dimensions in a pile, the work must be done expeditiously, and it is important that the matching of the feathers be perfect, inasmuch as the finished product depends on the symmentry of the feathers selected. It is important that the sorter possess excellent discrimination and the capacity to make swift decisions. That this test shows individual variations in ability to perform work of this order was proved conclusively by the correlations obtained.

The average time for the test was 255.33 seconds, with an average deviation of 80.63 seconds. There was a range in time of from 70 seconds to 360 seconds.

The sorting was on the basis of 100 per cent. for five neat piles of cards of similar sizes, irrespective of the time consumed in the sorting process.

No. one made less than five piles, but 35 per cent. made more than five. A system of penalties for extra piles was used, and each person making more than five piles had an additional five seconds for each extra pile added to her time-score.

The girls were arranged in order of merit on the basis of their standing on the Time-record.

The moderate workers did the best sorting, the swiftest made errors and their piles were untidy and slip-shod; the very deliberate, slow, painstaking workers saw differences where none existed.

III. Card Sorting.

A simpler test, that of arranging two packs of ordinary playing cards in suits was given to each one of the factory group.

Skill in this test depends much on practice but such psychological facts as neatness, system, or the reverse, may be noted, and the observations of this character made by the Experimenter in individual cases aroused the approval of the Foremen.

The clumsy girl let her cards fall frequently, or fumbled them. A lack of decision and keen perception caused loss of time and most of the errors. The sorting of cards is a simple eye-hand co-ordination exercise very like the work performed by the girls in the Selecting Department. The amount of discrimination necessary is reduced almost to a minimum, and in this regard it is more like the sorting of the Raw

Feathers than it is like the sorting of the finished products, the dyed and curled feathers.

The time ranged from 55 seconds to 181 seconds. The errors ranged from ten to one. By applying the Method of Penalties (46) described by Dr. Munsterberg fourteen seconds was added to each score expressed in seconds for every single error made. Upon the basis of the time-scores, the subjects were arranged in rank order of merit. The test was then correlated with the Criterion and the other tests.

This test helps to differentiate between the quick, neat, careful and resolute worker and the person who is vacillating, slip-shod and slow. It is distinctly an index of manual dexterity. The relation between Time and Score in this test shows an unusually high index, .75 P. E. .0397.

If an individual girl is being tested in order to determine into which of the Departments of the Factory she will fit best, this test is most illuminating. It throws light upon the girl's temperament and her type of performance reveals whether she will do good work as a Select or be better employed in some other department of the factory.

The median score in this test is represented by the time score, 134.46 seconds. Unless a girl performs the test within this time limit, she will not excell as a selector.

IV. Discrimination of Color Test.

The ostrich feathers are received in their natural condition, and the work of dyeing, curling and dressing is all done, at the factory.

After the first sorting of the raw material, the feathers are sent to the Dyeing Department where they are colored and then steamed. When these processes are finished, the feathers are again sorted and arranged so that they may be fashioned into the designs ordered.

There is sometimes a very slight difference in shade that is rather difficult to detect. An expert sorter must be able to differentiate between the slightest variations in color as well as in form and size.

In order to test the ability of the Selectors to do this, a Discrimination of Color Test was devised. Pieces of knitting yarn of various colors, fifteen inches long were tied in small bows. There were four sets of ten shades each of Blue, Purple, Orange and Green, which were mixed with seventy-five bits of yarn of other colors. Each set contained exactly the same number of pieces of yarn, namely, eighty-five small bows of yarn. These were placed in an envelope on the outside of each of which was written either the word ORANGE, PURPLE, BLUE or GREEN. In the envelope bearing one or the other of these color-names all the shades of the particular color specified were placed.

The directions were given orally thus: Each one of you has been given a sheet of paper. Write your name on the paper. You have been given an envelope also that contains bits of yarn of different colors. On each envelope is written the word ORANGE, GREEN, PURPLE or BLUE. That means that in each one of the envelopes there is a set of various shades of one or the other of these colors. Those of you who have an envelope marked "PURPLE" are to put all the shades of purple in a pile on the sheet of paper. Those of you who have an envelope marked "Blue" all the shades of blue, and so on in the same way with the "Orange" and the "Green." Pay attention to no other color than the one you are told to select.

Do you all understand? When I say "Ready," empty the contents of the envelope on your table and begin at once to sort your color. Remember to put the yarn you select on the white paper. When I say "Stop," do so at once. Sort the colors as quickly and as carefully as you can.

One minute was allowed and in that time one person had sorted all the colors. The other girls had selected from 42 per cent. to 93 per cent. of the colors.

Since it was desirable to ascertain those persons who were capable to making quick decisions coupled with a fair degree of accuracy, special stress was placed upon the time-element when the directions were given.

After a girl had sorted one set of colors she was given another set until each girl had had each set of colors to sort. This was done to see whether there was more difficulty in distinguishing between the shades of one color than another, and whether there was a gain in time through practice. The shades of purple seemed more easily and quickly selected inasmuch as 85 per cent. were picked out an average by the group in the allotted sixty seconds. The other colors in their apparent order of difficulty were green, orange and blue, with an average of 80 per cent.; 73 per cent.; and 60 per cent. selected within the time-limit.

When the results of the four color sortings were averaged the subjects were arranged in rank order according to the record made by each individual in the entire set of colors.

This ranking was then compared by the rating given to these girls by the members of the Firm.

V. Motor Control Test.

Most of the work in the factory is monotonous. The same type of movement is repeated over and over. The ability to stick to the task in spite of its extreme monotony, and create interest where none is apparent, is possessed by

some persons but is totally lacking in others. A girl, who because of a nervous temperament is unable to keep repeating the same form of activity, cannot do good work as a Bender or Finisher of Feathers; neither can she do her best work in the Selecting or in the Boa Department for as our Job-Analysis indicated, practically all of the workers, save those who run errands or are in the Fancy Department, must do the same thing over and over again.

In order to determine which of the employees are best suited for steady and monotonous labor, a new and very simple test was presented. On a sheet of 8 x 11 paper, 1088 oblique lines one-quarter of an inch long, of this character / / / / / were arranged in 29 rows, thirty-eight in a row. The test was described to the subject, the sheet given to her face down, and she was directed to make a cross out of each oblique line as quickly as she could. This was given to small groups though it may well be used as an individual test.

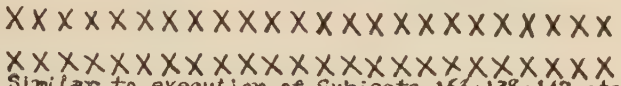
The subjects were told that the Experimenter would announce the thirty-second periods, and they were to check at what-ever point they happened to be working when the number was called. In this way, comparisons between subjects could be calculated, the effect of practice noted, and the individual differences in procedure observed. The length of time allowed for the test in optional with the experimenter. Five periods of 30 seconds each, or two minutes and a half was the allotted time for standardization in this case, although certain subjects were permitted to finish the sheet.

There was not as great a variation in the amount of work accomplished in two and a half minutes as there was in the quality of the work performed. (Five grades of achievement were arranged. See Page 57.

The medians for these subjects averaged 294.875 strokes.

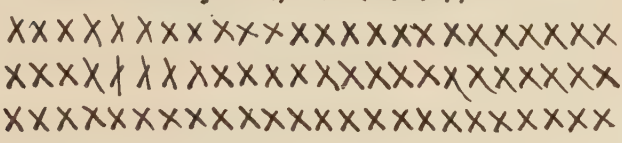
The general appearance of the sheet proved to be the most enlightening and important feature of the test.

Grade I. The most excellent type. Steady, even cross-lines maintained throughout. (90% and over).

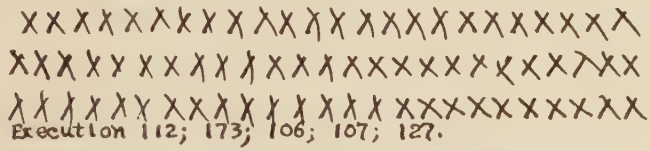


Similar to execution of Subjects 166; 138; 147 etc

Grade II. Even and Steady in spots. (80%-90%).

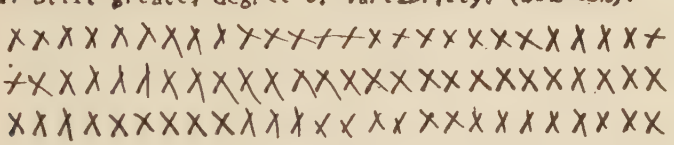


Grade III. More variability than Grade II. (65%-80%).



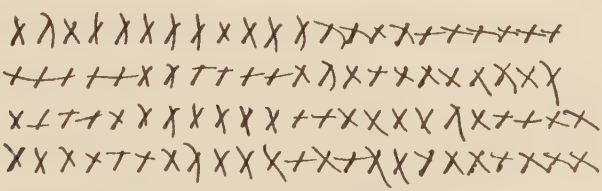
Execution 112; 173; 106; 107; 127.

Grade IV. Still greater degree of variability. (30%-65%).



Similar to execution of 158; 157; 142.

Grade V. The poorest execution. (Below 30%).



From the records of 119; 143; 151.

The papers were graded by three judges, foremen and directors in the factory. First, they graded the papers on the basis of Neatness; Second, according to the number of strokes made; Third, according to the length of the stroke, the shorter and more firmly the stroke was made, the higher was the grade given to the paper; Fourth, the uniformity of the cross stroke, when the strokes were approximately the same length above and below the oblique and when they crossed at about the center of the oblique they were rated higher than when the reverse was true. The Judges were asked to arrange the papers in rank order of merit and were asked to consider first one characteristic then the other in the different characteristics mentioned above. They were requested to pay no heed to any attribute other than the one specifically designated.

After each grading, the Experimenter noted on the back of the sheet the rank order given by each judge in each of the features postillated. The different ratings were then average, and a complete scale derived. This rarking was compared with the results obtained by the same employees in the tests as well as with the rank order on the Criterion Scale.

Personal traits of character are revealed by even a quick inspection of these test sheets, and a capacity for precision of movement can be noted as well as the subject's attitude toward monotony. Many of the girls began the test in a regular and steady manner but grew careless after a few lines were done and continued the task in a slip-shod fashion. Steadiness of motor control in these persons was so limited that the monotonous work of continually handling feathers palled on them and as the indices of correlation which are discussed, shows, this test is the best index of capacity to perform this particular grade of factory work that has been found.

The test met with the highest verbal commendation from the Firm. Irrespective of any mathematical calculations in connection with this test, the Steady Workers who tended to make rather even, well-balanced lines and the careless, erratic Workers who were inclined to change their methods of crossing the lines perhaps several times during the test, could be sorted out after a brief inspection.

E. Feather Sorting Test.

When the feathers are received, they are distributed in mixed bunches to the girls in the Raw Feather Department, who are directed to arrange them into piles. In this factory, the different types of feathers are designated as "Amazons," i. e. the broadest and widest feathers; the 18

inch," i. e. those 18 inches long; the "three fourths," i. e., those that are as broad as the "Amazons" but shorter; the "Byocks," which are from the tail of the ostrich and are black and white; the "Blacks," i. e. all black, as the name signifies; and "No. 1055," scrawny feathers, thinner and narrower than the others. The girls were told to lay these feathers into the six piles into which they were accustomed to place them. The Firm wished to have a report on the Time Rates and the Accuracy of the girls in this group. Such a test could only be of value when applied to practised. factory hands, and was to be used, in this instance to determine which persons should be laid off in the approaching slack period. For weeding out purposes, the test proved to have a distinct value.

Seven ounces of feathers were given as a standard amount: They were laid before the subject en masse. The time for sorting this amount varied from 167 seconds to 329 seconds, with an average time record of 227 seconds; an A. D. of 44.33 seconds; and a S. D. of 11.37 seconds.

There was a range in accuracy from no errors in one instance to fifteen in another. The average error was 4.83; the A. D. was 5.74; the S. D. was 7.29.

Those whose time-score was nearest that of the average, made the best score for accuracy. The very quick and the very slow, had noticeably poor records for accuracy.

A confusion of the Byocks with the Black Feathers, or a failure to discriminate between the other varieties counted as an error and girls were penalized 5 per cent. for each mistake which was deducted from the standard 100 per cent. perfect.

Although the feathers in other factories are occasionally designated differently, the time-records established by this fairly representative group forms a reliable basis for judging the ability of individuals engaged in this type of feather work.

F. Discrimination of Weight and Shape Test.

The girls in the Selecting Department were examined still further by another test designed to determine their quickness of perception and ability to discriminate between small differences in weight.

Eight wooden solids one one-half inches in diameter were so fashioned that there were four pairs, each pair of which had similar bases and sides. Thus there was a pair of seven-sided, eight, nine and ten-sided solids.

These blocks were weighted after the manner of the of the Binet-Simon weight Test—3 grams, 6. 9, 12. 15, 18, 21, and 24 grams, respectively. There are more blocks in this set than in the Binet, but the same difference between the weight of the blocks has been maintained, that is they were made progressively heavier by three grams.

Two distinct problems were given: (I) Put the blocks in a row according to their weight,—first, the heaviest, then the next heaviest, and so on. (II) Put blocks in pairs, according to the number of their sides. There are two each of like sides.

A time-limit of one minute for each part of the test was allowed. A record of each individual's arrangement of the weights was kept, two correct arrangements out of three were required for credit. Seventy-five per cent. received credit.

The nine and ten-sided figures were most often confused. Fifty per cent. of the girls performed the pairing in 45" without error. The other girls made from two to six errors each and required the entire minute allotted in order to arrive at a decision regarding the blocks with more than six sides. The methods employed by the subjects were enlightening as to the ability of a subject to observe small differences in material, and to arrive at a positive decision regarding those differences.

The girls were graded on the basis of the average rank obtained as a result of the two parts of this test.

Chapter V.

THE CORRELATIONS

The high value of correlations is well recognized and as Prof. E. L. Thorndike has said, "The only sound principle by which one can interpret the tests is determined from the results of the correlations." (57).

The process of obtaining an index of correlation involves the following operations:

First, each subject was ranked One, Two, Three, Four, etc., according to her standing in each individual test, as compared with the rating of her companions in the same test. In the first ten tests the standardized methods of scoring have been employed. The manner of scoring in the special tests devised by the Experimenter is described in the Appendix.

Those subjects who have a like score were ranked in order on the basis of the time required to perform that particular test. That is, if two girls received a rating of 80 per cent. in the Hard Directions Test, for instance, and one of these girls completed the test in 75" while the other girl took 76" to do the same test, the one who was quicker would be ranked higher than her more deliberate co-worker. In case of a "tie" for a given rank, (this rarely occurred), the "ties" were divided in such manner as to keep the total number of ranks equal in the series. If, for example, two subjects ranked 4th in a test, making an identical score in both time and degree of accuracy each was assigned the rank 4.5 in the series thus replacing 4 and 5.

Then, each girl having been given a rank order of merit in each test, a comparative calculation of the inter-relations between each test and every other test was derived by means of the Pearson Formula, in which r, or the coefficient of correlation equals

$$r = 1 - \frac{6 \times \text{Sum of } D^2}{n\,(n^2 - 1)}$$

This coefficient is so derived that when its value is Unity the two variables have perfect concomitance, and when its value is Zero three is an absence of relationship, so far as the conditions of the experimenter are concerned. It is evident that the reliability of this coefficient increases with the num-

ber of cases compared and also with the magnitude of the r obtained.

Seventy-five subjects were tested in this investigation. a sufficiently large number to demonstrate in a fairly reliable manner the value of the tests as indices of the efficiency of the employees.

The Probable Error was determined by means of the Pearson Formula. $.6745 \frac{1-r^2}{\sqrt{n}}$ Unless a correlation be at least twice as large as its probable error it lacks significance, and of course, to be a truly trustworthy index it should befour or five times as large.

The inter-correlations between the tests are indicated in the accompanying table:

But since the results obtained in the various tests administered by the experimenter can have no value until they have been compared with a fixed standard not open to question, it was decided to secure such a standard in the form of an Efficiency Roster which should be compiled from ratings made by the Firm. It will be apparent that a rank order arrangement of the employees made by task-masters who watched

The Tests.	Binet	Trabue	Hard. Direction	Easy Dir. #1	Easy Dir. #2	Av. 3-Direc.	Easy Opp.	Mixed Rels. 1	Mixed Rels. 2	Substi.	Can. A	Can. 2	Prov.	Motor	Cards	Sort-ing	Color	Feath.	Weight	Judg.	Criteria
Binet-Simon	1.	64	69	91	63	73	31	41	43	54	42	40	51	36	40	49					53
Trabue Completion		1.	63	42	60	75	38	61	59	43	25	36	40	47	47	36					62
Hard Directions			1.	75	64	70	32	53	51	40	38	32	57	56	44	44					53
Easy Directions #1				1.	74	76	34	59	56	37	52	42	50	58	40	44					68
Easy Directions #2					1.	63	58	51	35	32	55	31	42	33	36	26					71
Av. 3-Directions						1.	48	54	44	56	44	29	28	60	34	43					70
Easy Opposites							1.	63	42	47	36	42	43	34	58	49					65
Mixed Relations #1								1.	80	33	29	50	23	34	45	42					45
Mixed Relations #2									1.	48	54	47	43	49	44	32					39
Substitution										1.	31	48	50	38	42	40					60
Cancellation -A.											1.	44	28	43	33	35					42
Cancellation Digits												1.	42	31	28	30					69
Proverbs-Eng.-Ait.													1.	59	30	47					39
Motor Control														1.	42	54					80
Card-Sorting															1.	56					73
Sorting Test																1.					56
Color Discrim.																	1.	61	26	56	53
Feather Sorting																		1.	37	64	76
Weight-Form																			1.	33	22
Judgment																				1.	65
Criterion-Firm &c.																					

the performance of these workers for a period of months is a criterion not to be gainsaid. No worker who had been less than six months in the Factory was examined. In order to give more authority to this index, separate appreciations were asked of four different individuals, the Manager, the Foreman and two of the Forewomen.

The procedure was carried on in this manner. The Experimenter arranged four sets of seventy-five cards size 3 x 5 inches. On each card of each set was written the name and the identification number of each of the seventy-five girl who had been submitted to the tests. On the back of each card the Experimenter wrote:—1. General Intelligence. 2. Reliability. 3. Promotion. The four judges were each given a set of these cards and were directed to rank the girls whose names appeared thereon in order of merit; first, on the basis of his personal opinion of the General Intelligence of the girl. The judges were asked to disregard every other characteristic except the intelligence of the individual. The subsequent rankings to be made were not mentioned until each judge had completed his rating. The same directions were given when the other rankings were obtained. After each ranking, the numerical order in which each girl was placed by the judge was noted.

A composite ranking was then derived for each girl by calculating her average position in all four scales in each of three characteristics, Intelligence, Reliability, and rank order for Promotion. Correlations between this Composite or Firm Scale show positive indices raging from plus 28 to 80.

This scale froms the basis for the Teams of Tests to be developed for practical application in the Factory.

Rank Order	Card-Sorting	Substitution	Sorting Test	Cancellation--A	Cancellation of Digits	Motor Control	Criterion Scale
			First Quartile				
1	63	58	29	71	65	66	24
2	46	45	71	1	47	29	47
3	12	29	21	66	13	38	30
4	30	38	50	34	39	45	75
5	49	5	15	38	54	24	65
6	24	25	1	63	20	30	36
7	75	36	46	68	52	47	13
8	13	23	24	9	29	71	45
9	43	49	25	45	75	36	66
10	65	65	66	16	26	65	21
11	68	26	7	3	6	58	32
12	62	66	62	65	1	62	48
13	40	14	65	39	30	43	50
14	36	47	30	42	9	31	25
15	21	24	63	61	48	50	9
16	56	39	13	70	4	32	15
17	47	63	59	36	40	12	59
18	8	71	39	53	66	15	64

Rank Order	Card-Sorting	Substitution	Sorting Test	Cancellation-A	Cancellation of Digits	Motor Control	Criterion Scale
			Second Quartile				
19	45	59	34	29	58	5	5
20	15	30	74	30	32	48	63
21	32	35	16	5	21	25	22
22	9	70	40	10	59	59	58
23	64	32	47	24	5	3	53
24	70	37	68	48	71	22	64
25	50	75	31	54	24	34	16
26	66	72	12	15	50	9	39
27	31	21	75	40	25	21	6
28	22	48	32	59	42	23	37
29	5	31	38	37	43	75	17
30	59	7	72	64	62	5	71
31	37	42	10	50	74	41	72
32	39	52	45	58	17	64	31
33	25	58	28	75	15	68	26
34	71	40	7	43	26	26	7
35	58	53	5	8	68	39	67
36	6	6	37	20	8	72	29
			Third Quartile				
37	60	46	22	23	45	52	46
38	1	67	35	46	63	49	40
39	14	15	3	47	46	16	1
40	17	18	64	55	10	1	54
41	26	16	70	57	12	70	23
42	67	4	58	12	3	14	62
43	41	12	8	17	37	33	35
44	16	3	4	6	50	73	38
45	44	50	42	62	36	46	44
46	3	22	36	4	28	53	60
47	10	1	67	13	41	27	21
48	20	28	52	22	61	11	70
49	23	2	60	33	23	18	74
50	29	8	56	11	19	13	52
51	34	74	41	21	67	63	12
52	38	60	73	52	22	74	49
53	42	10	61	19	38	57	61
54	52	69	53	2	33	42	68
			Fourth Quartile				
55	35	62	11	32	60	8	8
56	27	17	54	25	31	6	10
57	18	19	2	49	18	67	4
58	4	55	43	14	16	69	55
59	28	56	44	60	11	7	11
60	53	9	48	18	70	50	56
61	2	41	69	31	73	60	69
62	72	61	26	51	27	20	48
63	54	64	17	72	7	10	2
64	55	13	55	67	28	35	73
65	19	43	49	73	72	2	42
66	23	27	18	26	69	56	34
67	57	44	19	27	14	7	57
68	7	68	27	56	55	61	27
69	51	57	23	28	35	44	41
70	73	20	20	41	64	28	28
71	48	34	14	44	49	55	19
72	74	73	6	35	44	4	3
73	11	54	33	69	53	37	18
74	20	11	57	74	51	19	33
75	69	51	51	7	57	51	51

Rank Order	Easy Opposites	Mixed Relations No. 1	Mixed Relations No. 2	Trabue Compl.	Proverbs Test	Easy Directions No. 1	Easy Directions No. 2	Hard Directions	Av. 3 Directions	Binet Simon	Rank Order Binet
					First Quartile						
1	59	1	29	39	5	24	71	58	65	65	1.5
2	24	30	70	59	65	36	75	36	75	48	1.5
3	75	24	75	48	39	15	59	75	71	75	5.5
4	65	70	63	5	16	75	15	59	24	45	5.5
5	29	65	38	65	71	59	6	39	36	47	5.5
6	1	39	52	75	66	65	59	70	34	24	5.5
7	15	45	59	58	54	15	65	24	30	71	5.5
8	36	39	7	29	58	63	37	40	5	1	5.5
9	14	16	39	30	74	45	14	36	47	59	10.5
10	71	5	24	17	47	58	48	43	6	29	10.5
11	45	38	66	70	61	47	36	63	15	70	10.5
12	26	16	71	45	24	48	5	5	48	24	10.5
13	58	9	60	21	14	66	21	38	21	25	15.
14	17	29	30	32	34	21	39	52	45	20	15.
15	48	42	1	22	63	6	64	30	58	31	15.
16	70	71	36	61	52	20	7	6	9	21	15.
17	45	40	20	40	43	17	45	48	64	32	15.
18	68	36	16	6	1	37	52	21	7	50	15.
					Second Quartile						
19	30	61	34	36	72	64	46	14	52	6	15.
20	52	59	55	66	38	14	72	29	40	58	15.
21	38	72	61	50	64	46	1	10	66	37	15.
22	20	12	42	52	45	7	49	64	29	46	15.
23	12	60	45	14	17	29	8	15	38	49	15.
24	53	6	37	15	53	54	12	49	1	72	25.
25	32	68	9	71	40	39	55	66	49	62	25.
26	21	21	13	34	36	72	23	74	16	43	25.
27	50	56	72	1	20	40	61	22	46	23	25.
28	63	52	53	67	13	22	58	33	32	9	25.
29	37	74	58	16	8	23	9	20	12	63	25.
30	9	23	6	54	60	50	61	17	23	15	25.
31	72	48	21	60	9	16	66	50	20	56	25.
32	23	28	23	33	48	71	74	7	34	39	25.
33	13	53	17	26	23	12	40	53	42	22	38.5
34	63	41	12	42	26	61	42	34	10	11	38.5
35	31	8	13	46	12	32	4	52	22	15	38.5
33	33	53	17	26	23	12	40	53	42	22	38.5
34	63	41	12	42	26	61	42	34	10	11	38.5
35	31	8	13	46	12	32	4	42	22	54	38.5
36	60	17	28	27	25	31	43	23	74	13	38.5

Rank Order	Easy Opposites	Mixed Relations No. 1	Mixed Relations No. 2	Trabue Comp.	Proverbs Test	Easy Directions No. 1	Easy Directions No. 2	Hard Directions	Av. 3 Directions	Binet Simon	Rank Order Binet
					Third Quartile						
37	35	15	6	28	31	38	50	1	43	42	38.5
38	7	26	50	56	10	8	53	4	61	31	38.5
39	15	22	56	38	68	1	13	37	25	54	38.5
40	34	31	27	8	21	33	19	28	60	52	38.5
41	74	63	43	10	37	28	33	46	72	7	38.5
42	4	7	19	73	56	41	38	54	8	4	38.5
43	25	18	46	43	22	74	32	60	4	55	38.5
44	64	10	73	23	41	11	18	72	26	3	49.
45	42	54	49	12	62	13	28	26	33	40	49.
46	44	3	11	13	28	34	60	62	31	14	49.
47	43	4	15	49	27	67	31	13	56	41	49.
48	11	55	18	57	42	56	50	67	13	27	49.
49	2	19	54	41	35	19	17	61	18	26	49.
50	67	11	74	3	19	26	68	56	67	35	49.
51	41	73	3	68	57	35	41	19	27	12	49.
52	19	32	37	11	3	42	2	8	27	68	49.
53	28	46	41	2	73	27	44	57	55	3	49.
54	57	49	67	4	15	73	10	55	11	28	56.
					Fouth Quartile						
55	8	13	14	53	34	68	74	44	41	60	56.
56	61	64	25	20	32	18	56	73	68	10	59.
57	3	67	4	74	32	55	25	41	73	68	56.
58	18	2	48	35	4	44	51	11	2	56	56.
59	70	71	20	61	43	22	7	6	29	51	59.
60	56	69	27	25	54	62	35	69	42	19	59.
61	73	33	4	44	18	3	57	27	3	57	59.
62	69	57	2	69	51	69	3	35	51	18	59.
63	38	72	61	50	64	46	26	14	37	49	59.
64	10	14	35	55	46	4	67	31	2	28	68.
65	40	54	41	12	62	13	28	62	33	43	68.
66	8	13	14	50	44	68	74	41	19	60	68.
67	35	10	6	28	31	38	53	4	61	61	68.
68	71	55	44	62	60	57	11	2	57	44	68.
69	73	30	4	44	18	3	69	30	51	2	69.
70	19	32	31	11	73	42	2	8	27	33	68.
71	5	20	69	25	54	62	62	69	46	69	68.
72	6	25	28	62	69	57	35	2	57	44	72.
73	16	27	33	44	18	3	57	27	3	57	73.
74	22	34	57	51	55	69	3	28	51	18	74.
75	27	51	20	69	44	33	20	51	69	51	75.

When the members of the Firm were asked to arrange all the employees in rank order on the basis of their suitability for Promotion, they stated that some persons would not be promoted under any consideration. All four judges agreed that twenty-four out of the seventy-five employees would remain statu quo. The twenty cards representing these

persons were removed from the pack. The Manager then consented to arrange this collection of non-promotable persons in rank order on the ground of their general efficiency. According to this ranking, the twenty-four employees designated were given grades after the fifty-one had been placed in order for Promotion. The Judges arranged them also in rank order on the basis of the other two characteristics and the average position thus derived placed them in gradation with their confreres.

It is apparent that the four judges may disagree as to the exact rank of the individual girl; their ideas as to what is meant by the terms General Intelligence and Reliability may be unlike, but the average position achieved by each girl in these three characteristics cannot be biassed by individual misapprehensions. Ranked order scales in estimating the efficiency of a large group of persons are admittedly fairer when a combined rating of several competent judges is made.

By an inspection of the Table of Correlations, certain facts of interest may be noted.

The easiest tests bear a higher relation to the Criterion than do the more difficult tests such as the Proverbs and the two Mixed Relations Tests.

The results in the Mixed Relations Tests do not confirm the findings of Wyatt (51) who states that they form the highest correlation with Intelligence of any of the tests save the Trabue Completion.

The Coefficients of Correlation between the Special Tests and the Criterion demonstrate that the Card Sorting, the Sorting Test No. 2 and the Motor Control Test are of value, and inasmuch as they represent motor activities similar to tasks performed in the Factory they should be regarded with special significance. The Card Sorting Test gives an index of plus .73., P. E. .0280. The Sorting Test No. 2 a positive index of .60., P. E. .0498. The Motor Control Test gives the highest coefficient of any of the tests that were presented to the group, i. e. plus .80., P. E. .02484.

Both of the Mixed Relationship Tests show fairly high indices of correlations between them and the Directions Tests, and between them and the Linguistic Tests. In studying the Table the low inter-correlations between these two tests and the purely mechanical tests are not surprising. In the records of individual girls it is apparent that with a few exceptions some of them do consistently well in all tests that demand on or the other type of ability, scholastic or motor.

The Cancellation Tests correlated with each of the other tests give low but positive indices. As was noted under the discussion of the Binet Test, this type of test does not bear a high correlation with tests that demand General Intelligence.

Dr. Weidensall found a positive index of .52., P. E. .055. between the A-Test and General Intelligence (31), while our coefficient is but .36., P. E. .0681. In the same group of Bedford women tested, she found a correspondence in this test between the rate of its performance and the difference in intelligence as indicated by school grade. This was to be noted in the Factory group also. By a comparison made between the subjects arranged in order on the basis of their time-records in the A-Test and the grade in school attained at the age of fourteen years, one finds the retarded pupils required 7 per cent. longer than did the median of the group to perform the test.

Terman and Bagley obtained a negative correlation and Simpson (59) a slightly positive one between intelligence and skill in the A-Test.

In "School and Society" (Vol. V., P. 24), there is a report of an experiment in which the Cancellation Tests correlate negatively with the Composite Tests which are adjudged good measures of mental acumen. To quote the statement of Dr. McCall, "This proves that a negative correlation may exist between apparently desirable traits."

In this study, however, the Cancellation Tests bear positive rather than negative relations to the other tests and would seem to uphold Prof. Thorndike's (60) theory that there is not an inverse ratio between desirable traits.

Correlated with each other, the Cancellation A-Test and the Digits Test indicate a relationship of plus .44., P. E. .0629. This is not, comparatively speaking, a high correlation for between tests that demand the same type of reaction there is usually a closer index of relationship. For instance, between the two Mixed Relationship Tests there is a positive coefficient of .80., P. E. .0280., and between the three Directions tests there are fairly high indices of correlation.

The reason for this case of low inter-correlation between two tests of a similar type may have been a lack of interest in continuing at work of the same kind, evidenced by several persons when performing the second cancellation test.

From an inspection of this table, there are no indications of such a hierarchy of coefficients as is demonstrated by Burt's "Hierarchy of the Specific Intelligences".

Ranked in the order of the degree in which they correlate with the Criterion, the tests are Motor Control, the Card Sorting test; Easy Directions No. 2, Averaged 3-Directions, Easy Directions No. 1; Easy Opposites; Trabue Completion, (1), the Substitution Test, Sorting Test, Hard Directions, Binet Proverbs, Digits, Average of 2-Cancellation Tests, Mixed Relations No. 2., A-Test.

Although there is no very special conclusion to be drawn in regard to the possession of "general ability", the more intelligent of the subjects tend to do well in whatsoever test they are given to execute.

The Special Tests.

Twenty girls chosen at random from the Selecting Department were submitted to the following tests especially devised by the Experimenter to measure their capacity to perform the tasks demanded in this department: (1). Weight and Form; (2). Color Discrimination; (3). Feather Sorting; (4). Judgment. These tests are described on Pages 50-60.

Before proceeding with the tests, the Manager, the Foreman and the Forewoman in charge of the Selecting Department made a rank order arrangement of this particular group of employees on the basis of the "Efficiency displayed by each in their daily work". The average of these three ratings formed a new scale, the Criterion, with which the girls' standing in the four tests have been correlated.

Rank Order of Selecting Department Girls in Tests indicated. Criterion—Three-fold estimate of Firm.

Rank	Color	Judgment	Weights	Feathers	Criterion
1.	21	28	63	24	63
2.	53	24	33	74	24
3.	28	46	28	49	28
4.	24	63	21	63	68
5.	74	68	12	21	64
6.	63	53	49	68	21
7.	50	21	64	28	50
8.	12	64	8	46	46
9.	49	12	74	8	74
10.	64	8	41	53	1
11.	8	49	71	1	12
12.	46	74	50	41	49
13.	71	33	68	12	16
14.	16	71	53	50	53
15.	69	51	24	64	69
16.	41	1	69	33	41
17.	68	16	1	69	8
18.	1	69	46	16	51
19.	33	50	51	71	71
20.	51	41	16	51	33

a The coefficients were derived by means of the Formula

$$r = 1 - \frac{6 \times \text{sum of } D^2}{n\,(n^2 - 1)}$$ In this study $n = 20$.

The result of this correlation is shown in this table:

	Wt. Form	Color Dis.	Feathers	Judgment	Criterion
Wt. Form	1.00	.26	.37	.33	.22
Color Dis.		1.00	.61	.56	.58
Feathers			1.00	.64	.76
Judgment				1.00	.65
Criterion					

The Feather Sorting Test which involves the use of the material the girls are familiar with gives the highest coefficient of correlation, .76.

The Judgment Test which the Firm found of so much value that it was adopted immediately by the Employment Forman, ranks second. This test measures the cacacity for making quick and accurate decisions, and is also similar in its demands to the work of the department.

Time vs. Accuracy.

One of the differences between a superior and inferior worker lies in the speed with which each performs his task.

In studying the individual variations of the members of this group two fairly distinct types of time-scores may be noted the time-record of the slow, methodical worker and that of the nervous, quick, distractable worker. The deliberate and painstaking girl stood higher in the estimation of her employers than did the quicker and more alert girl who was inclined to be inaccurate and careless. The latter made good time records but in excellence of performance, her score was low. Many of these girls over-emphasized the desirability for haste and rushed through the tests, assuming that the faster they worked the better it would be for them. This was done regardless of the admonition to "work carefully". Such reactions were noticeable in the first written tests performed by Subjects No. 25, No. 56, 60, and No. 70. in the Easy Opposites and in the Mixed Relations Tests. Speed and Accuracy with them revealed a negative correlation.

The reactions of the employees numbered 65, 75, 24, 71, 15, and 63 for example, showed a fair and consistently high rating in the tests so far as accuracy or performance went but their time-scores were variable and not below average length.

The Coefficient of Correlation between Time and score for this group is usually highest in those tests in which a

demand is made primarily upon motor activity. Thus, in the following tests there is a positive correlation between time and score of:

Test	Coefficient	P. E.
Substitution	.44	.028.
Digits	.57	.0361.
Sorting Test	.62	.031.
Motor Control	.65	.019.

In the Linguistic Tests, the indices between Time and Score show positive but low correlation, thus:

Trabue Completion—.23., P.E. .081.; Proverbs Test, .20., P.E. . . .015.

The Directions Tests give positive indices of .50., 51. and .53., respectively.

Between Time and Score in the Mixed Relations Tests there was a positive coefficient of .27., P.E. .037., and .31., P.E. .036.

While it cannot be considered conclusive evidence that the correlation between Speed and Accuracy is always higher in tests involving motor activity when such a group as this is examined, the probability is that this would be the case.

The question arises as to the value that should be placed upon Speed alone as a legitimate index of mental efficiency. In the Factory, Speed per se has, in general, a negligible appraisal. Unless Speed is accompanied by Accuracy, the Firm announced it an undesirable trait in an employee.

It is no doubt true that when some individuals, working under constant conditions, increase their speed, they tend to increase the number of errors, but other individuals under like conditions may show a decrease in their errors. Prof. Whipple (7) quotes results that show that the faster subjects are also the most accurate. In six of his groups, Brown (58) found positive correlations between Speed and Accuracy of Adding that ranged from .13 to 43., with P. E.s ranging from .07 to .12. With small groups of college students, Prof. Whipple states that he has obtained positive correlations of .19 in the case of adding and as high as .86 for mental multiplication. (7).

The Factory Girls cannot be judged by the same standards as the College Students in the reaction time required to perform mental tests.

All save those subjects whose standing in the tests placed them in the upper third of the group react slowly to all stimuli. Their cerebrations are deliberate even when the simplest matters are under consideration. In some cases, the slow reactions are due to the fact that the subjects are unaccustomed to writing, and the co-ordinations necessary for

this form of activity show disuse. Though the median age of the subjects was seventeen, there are several of the older women have been out of school a great many years. For them the task of writing their responses involved an actually laborious application of unhabitual motivity.

The weight, then, that is to be placed upon an individual girl's time-record in her performance of tests in this Factory must depend upon her accuracy of achievement rather than upon the time element alone.

Chapter VI.

1. MULTIPLE CORRELATIONS.

The correlations obtained between the results of the employees' standing in each of the several tests and the ranking of these employees made by the Firm vary according to the particular test applied. However, in the sum-total of reactions they indicate that the verdict secured through the tests approximates the valuation established by the Firm especially as it affects the upper and lower quartiles of the Scales involved.

Meumann (59) states that it is an error frequently made by the exponents of what he calls "correlation psychology" to conclude that because a high correlation exists between any two elements it is necessarily a proof of the good qualities of both. "Both", he states, "may be one-sided". A combination, therefore, of several tests should be made in order to obtain their composite valuation. Thus any criticism that too great emphasis has been placed upon the result of the correlations between the Criterion and single tests may be forestalled.

Stern (60) and Burt (61) found that the higher correlation of the combined results of several tests show that one test compensates for another, each calling forth particular mental functions, so that only through such combination can a total picture of intelligence be created. The combined coefficients obtained from several tests have been calculated, therefore, in this group.

Indeed, one of the objects of this investigation is to demonstrate the combinations of tests that may be given practical application at the Factory. For it is an obvious fact that the index derived from the presentation of a group of tests will be a more reliable indicator than will the single isolated test that is naturally more or less narrow in scope.

The selection of tests to form a Team should be made on the basis of the magnitude of the coefficients of correlation.

"A Multiple Correlation Coefficient is that correlation which expresses the total efficiency of the scale when the tests choosen are those that bear the best or highest correlation with the Criterion." By this method there may be constructed a combination of tests which will predict "general ability" better than can be done by a single test.

Tests for a Team should bear high correlation to the Criterion and low indices of correlation with each other.

High inter-correlations indicate the existence of a too-close resemblance between the tests to make their combination and subsequent use as a Team worth while.

Recently, Dr. Herbert Toops of the Institute for Educational Research has devised a new method for the more rapid calculation of Test-Combinations than those formerly in use. He has transmuted algebraically the old method of procedure and by means of a printed form made it possible to reduce the formula, $\frac{(ric)(riu) \quad 2(ruc)(ric)(riu)}{1-(ruc)^2}$ to a series of more or less mechanical operations. This method has not yet been published, and I am indebted to him for his verbal demonstration that has already facilitated this portion of the present study.

In general terms, the rule is to select from the Table of Correlations that test which bears the highest relationship to the Criterion. Ordinarily, the subsequent choice of a test should be one which bears a high correlation to the Criterion but a low index to the first test chosen.

We assume a weighting of 1.00 for the first test as its Coefficient more closely approximates unity than do any of the other tests. The object in selecting a Team of Tests is to determine that series of tests which will approach most nearly to the ideal relationship of unity.

Following the prescribed method of procedure, test after test may be added to the original pair and an increasingly higher coefficient be obtained without changing the weights of the tests already in the combination.

Occasionally, the gain in the increment added is not sufficiently large to make the particular test combined of value. When this occurs, that especial test is to be eliminated so far as that test-combination is concerned and another test may be chosen in its place. Herein lies one of the apparent advantages offered by this method.

As an illustration, in this study, the Easy Directions Test No. 2 with its positive coefficient of .71 was selected as the Key Test; to it was added the Substitution Test, with an index of .60 with the Criterion and its inter-correlation of .32. This combination gives an index of .8368. Next, the Digits Test was added with its Criterion Coefficient of .49., and its inter-relating coefficients of .31 and .42 which again raises the multiple coefficient to .8432. By adding the Sorting Test (Criterion Coef. equals .56., inter-correlations equal .30., 36.) to this combination, the coefficient becomes .8436., a gain of only .0013. Hence, it is advisable to drop the Sorting Test

in this particular combination, and in its place, add the Cancellation of A-Test which has a Criterion coefficient of .39. The Multiple Coefficient then becomes .8747. As these tests require as a maximum about ten minutes for their execution, the Team seems an extremely practical one.

The next combination selected consists of the Card Sorting Test with its coefficient of .73 with the Criterion and the Average of the Three Directions Tests which correlates with the Criterion with a coefficient of .70. The inter-correlation between the two is .55. Combined, the two tests give a coefficient of .8128. Now if we add to this combination the Trabue with its coefficient of .62 when correlated with the Criterion, and of .75 correlatd with the Three Directions, and of .47 correlated with the Sorting Test, we get an increase in our coefficient of .0049; to this combination, add the Digits Test (Criterion Correlation equals .49; Inter-correlations of .35., 42., .43.,) which raises the coefficient to .8217; to this we now choose to add the Substitution Test which has a coefficient of .60 when correlated with the Criterion, and of .56., .48., .43 and .40 when correlated with the rest of tests in the combination. The coefficient resulting from this operation rises more decisively and gives us a coefficient of .8507

Judging from the time-records made by this group of girls, this entire set would require approximately 22'48.5". Some of the other test-combinations have proved of greater value because less time is required for their execution, and they offer a larger variety of mental exercises.

For instance, the Motor Control Test which combined with the Sorting Test (Criterion Coefficients of .80 and 73; inter-correlation of.54) gives an index of .8667. Adding to it the Easy Directions No. 2, (Criterion Index equals .71., inter-correlations equal .36: .33), and the Multiple Coefficient becomes .8773., an increase of .0106. If the Easy Opposites be added, (Criterion Coefficient equals .65; inter-correlations equal .49., 34., .68), the Coefficient becomes .9292. This is a good practical combination inasmuch as it may be administered within a limit of 500", and combines variety and brevity.

Another combination consists of the Easy Directions No. 2 and the Card Sorting Tests (Criterion Coefficients of .71 and .56 and an inter-correlation of .34) which gives a coefficient of .7867. To this, add the Digits Test (Criterion Coefficient equals .49; inter-correlations of .33 and .31) and the Multiple becomes .8118. Then add the Substitution Test (Criterion Coefficient equals .60; inter-correlations equal .32; .43; .48) and our Multiple Index becomes .8409. These tests require about 675." and have already proved of value in practical application at the Factory.

The next group is comprised of the Card Sorting and Three Directions Tests Combined, (Criterion Coefficients

equal .70 and .73 and inter-correlations equal .58); together these two tests give a coefficient of .7812. When the Easy Opposites Test is added with its Criterion Coefficient of .65., and inter-correlations of .58 and .45, the coefficient becomes .8327. Digits added to the series increases the coefficient to .8544. This combination requires about 807″ for its execution.

Another combination is comprised of the Motor Control and the Easy Directions No. 2, which give a Multiple of .9071; Sorting adds .0494; Substitution .0160; and Digits .0027: making a total coefficient of .9681., the best combination that has been derived. It is composed of tests which can be given in approximately 13 minutes. Each one of these tests has been given a practical trial by the Employment Managers and they are continuing to prove their worth as a rapid means of predicting the efficiency of an employee.

The Special Tests devised for use in this particular industry were next combined by taking first the Feather Sorting the Judgment Tests with their Criterion Coefficients of .76 and .65 and their inter-correlation coefficients of .64, whereby the Multiple Coefficient becomes .7944. When the Color Test is added to this group the coefficient becomes .8002. The Weight Test make no contribution of sufficient size to warrant the time required for its presentation.

It is to be noted that in determining these Teams, only tests bearing the highest indices of correlation with the Criterion have been considered. Thus, the two Mixed Relations, the Proverbs and the Weight Tests have been dropped. So far as this type of work is concerned they do not predict failure or success, their contribution is practically nil. The Binet requires too much time to be practical; only the single tests that have been designated elsewhere are to be retained.

To ascertain the weighting of the tests retained, the tables and procedure arranged by Dr. Toops were applied.

The first step is to choose that test which bears the highest correlation to the Criterion. Each test in turn is combined with the first test to ascertain what the special contribution of each to the combination already determined may be.

We take, then, the Motor Test as the "combination" when the Binet is combined with it, the Binet has a weighting of .6736; the Motor and the other tests have a weighting as follows:

Motor and Trabue	give	a	weighting	of	.4797.
Hard D.	"	"	"	"	.2778.
Easy D. No. 2	"	"	"	"	.7398.
3-D. Av.	"	"	"	"	.4146.
Easy Opp.	"	"	"	"	.2921.

Mixed Rels. No. 1	"	"	"	.2751.
Mixed Rels. No. 2	"	"	"	.0003.
Substitution	"	"	"	.5175.
A-Test Cancel.	"	"	"	.1347.
Digits	"	"	"	.3735.
Proverbs	"	"	"	.1439.
Cards	"	"	"	.7137.
Sorting	"	"	"	.2572.

By inspection the Easy Directions No. 2 and the Cards Test have the highest weighting. The first two combined with the Motor Test make a contribution of .8968,; while Cards and Motor have a correlation of .8666. Substitution with Motor as a combination have a correlation of .8616; the contribution of any of the other test is of less value at this point.

Taking Motor and the Easy Directions No. 2 as the combination, on the basis of the weightings already determined, we add the Trabue, which at this point receives a weighting of .1880; the Hard Directions has a weighting in this combination of 1419; Easy Directions No. 1 of .0169; Easy Opposites of .1201; Mixed Relations No. 1 has a weighting .0321; Mixed Relations No. 2, a weighting of 1363; Subsitution has a weighting of .5162; the A-Test of weighting of .1828; and Digits has a weighting of .3348; Card-Sorting, a weighting of .9390; Sorting Test a weighting of .1767; The highest of these weightings, the Card-Sorting Test, is, selected for the next combination.

With this new combination, we then derive the weighting of each of the other tests, in turn. Thus, when Motor, Easy Directions No. 2 and the Card tests are combined, the Trabue contribues .0823; Substitution .4734; Digits .3382. The other weightings are of less value. When Substitution and Digits are added to the former Combination, the aggregate correlation is .9591. Each test makes some contribution to the whole weighting, but in this case, additional tests increase the sum-total contribution in only a slight measure.

The Special Test should now be considered. Feathers and the Color Tests combined give a weighting of .2866; Weights test has a negative value of .0903., and is, therefore, to be dropped. When Judgment is combined with Feathers and Color, the Feather test remaining as the constant, the weighting becomes .4756. The correlation then becomes .7893. With Color added, the correlation becomes .8002.

Surveying all the weightings in this group of test, the contributions made by the Easy Directions No. 2., the Substitution, Cards and Digits tests, if the Motor Test is chosen as a key test, have the greatest value.

Of the Special Test, the Feather Sorting and the Judgment Tests are to be selected in preference to the Color and Weights tests.

It is possible, therefore, by this method to determine the actual contribution made by each test, and let the weighting thus derived be a guide as to the retention or rejection of the tests under consideration.

Before practical use of these methods can be made, however, it is necessary to predict the scores within the limits of which the test-records of the subject must lie.

The first step in this procedure is to select those teams of test which bear the highest coefficient of correlation to the Criterion and contain the individual tests which prove of the greatest value in their practical application.

It is essential to choose more than one team because of the varied demands of the Factory work. Hence, teams No. 2., No. 3., No. 6 and No. 7. are chosen.

The Teams of Tests which are to be applied in this industry.*

Team II.

Trabue Completion
Sorting
Three-Directions
Digits
Substitution

Team III.

Motor Control
Sorting
Easy Directions No. 2
Easy Opposites

Team VI.

Motor Control
Easy Directions No. 2
Substitution
Sorting
Digits

Team VII. Special Tests

Feather Sorting
Judgment
Color Discrimination

It is of practical value to have a definite method for scoring each Team of Tests. For this purpose it is necessary to obtain the combined predicted score for each Team. This is determined by dividing the Beta Weighting (obtained by the method demonstrated above, of each test by the Standard Deviation derived in that test, and multiplying each individual's test-score by the quotient thus obtained. This resulting quotient is usually an unwieldy number, a decimal that is difficult to manipulate. It is, therefore, advisable to reduce all the decimals in the Team to a workable form. This can be done by either dividing or multiplying each quotient by the same number. The algebraic theorem which states that multiplying or dividing all the members of an expression by the same number does not alter the existing relationship

*The directions for giving the special tests are found in the Appendix

between the members, justifies this proceeding. Having determined by which process to reduce the quotients to numbers that will make convenient multipliers the next step in the procedure is to multiply each girl's test-score by the multiplier derived for each individual test. The sums of all the products is the total score obtained by that individual in the tests of that particular Team.

The range of test-scores varies. The significant value of this operation lies, therefore, in the fact that, all the test-scores are reduced to one level and a subject is said to be rated in a certain decile in the combination of tests, rather than in one decile in one test and in another decile in another test etc.

Now, in order to determine what may be the level of achievment demanded in Team No. 2, for instance, it is to be noted that since the possible scores range from approximately 4224 to 1418 as a combination record, these are to be regarded as test-limits. The accomanying table may be used as a sample illustration of the method.

The girls whose individual records range from approximately 3000 to the upper limit in Team No. 2 do excellent work; those whose records range from 2000 to 3000 are fairly good, while those below the 2000 limit are doing poor work. Therefore, we are justified in demanding a combination score in Team Number Two of approximately 55 per cent. of the highest score made.

The same method of procedure is to be followed in each of the other Teams.

Thus, the highest combination score for Team 3 is approximately 5000 and its lowest limit is about 1300. We find no subjects on this scale who are doing good work, whose records are below about 2500, which is approximately 60 per cent. of the average record of the highest Group in this test. In Team 3 the multiples are 19 for the Motor Test; 12 for the Sorting Test; 8 for Easy Directions No. 2 and 7 for the Easy Opposites.

In Team 6 the multiple for the Motor Test is 19; for the Easy Directions Test is 14; for the Card-Sorting Test is 9; for the Substitution Test is 3; and for the Digits Test is 1.

In Team 6 there is a range of possible scores of from approximately 1500 to 4000. In this case we shall call 60 per cent. of the average score of those in the highest quartile of the test a good passing mark. Thus, those whose record scores are below an approximate rating of 2500 will not be valuable assets in the Selecting Department.

The Special Tests, of which only 3 are retained, the Feather Sorting, Judgment and Color Tests, have these multiples respectively, 4, 6 and 3.

The range in this test is from approximately 1200 to 700 as limiting scores.

In this instance, those obtaining a combination score of approximately 1100 and over are to be rated as excellent, those whose scores range from 900 to 1100 are fair, and any person grading below the level of approximately 900 are to be ranked as poor. This group is to be given a 75 per cent. pass mark, or an obligatory score of approximately 900 points.

A record blank can readily be arranged upon this basis for each team of tests. . .

FORM TO BE USED FOR RATING SUBJECTS

Name	Fitness Credits				Total
	Test 1	Test 2	Test 3	Test 4	Scores
Maximum Scores					
	100%	97%	90%	100%	

PERCENT RATING OF 5 SUBJECTS IN TESTS OF TEAM VII.

No.	x 4 Feather Test	x 6 Judgment	x 3 Color	Total
24	100	95	86	1228
49	95	80	80	1100
12	75	82	83	1041
50	70	72	82	958
51	25	74	60	724

Chapter VII.

INTERPRETATION OF THE DATA.

No form of measurement, whether it be physical or mental can be absolutely exact, but the more tests we give, the closer is our approach to a perfect index of the subject's mental status. The large number of tests applied in this investigation was made because of a desire to obtain as complete a mental picture of the group as was possible.

The art of evaluating tests scientifically consists in making checks and counter-checks that it may be definitely determined to what extent the test-results may be depended upon. This has been done in our study by the methods of correlation. The results demonstrate (1) which tests are of greatest value from the view-point of the employer; (2) how tests compare in difficulty of performance; (3) what combinations of tests reveal the best standards by which to judge an employee's capacity for work in the factory; (4) the relative standing of individual girls; (5) the relative standing of the groups engaged in the different departments of the factory; (6) the indices of test-values derived from comparing the rating of the girls in the tests with their rating by the Firm as revealed in their daily work.

The aim has been to simply register all the positive indications of intelligence. By means of the test-results, different levels of proficiency requisite for a person who is to be engaged in any one of the different departments of this particular type of factory have been established. The "levels" vary with the demands of the several departments.

In the second chapter of this investigation, the work required in each of the departments of the factory is analyzed. This study would be incomplete if the analysis of the requirements of the job and of the proficiency of the girls as established by the tests were not brought together. Subsequent chapters have been devoted to the analysis of test-records and a comparison of those records with the rankings of the firm.

The results of the application of methods to determine the Partial and Multiple Correlations established between the tests and the Criterion are in themselves sufficiently significantto require no further comment regarding their value. The "Teams of Tests" chosen by these methods of selection

and elimination demonstrate the best means by which to secure a swift and accurate picture of the applicant's native capacity, in so far as that capacity will prove of use in this type of work.

In the Fancy Department, it has been demonstrated (Page 42) that the ability to reproduce the two designs in cluded in the Binet Scale is a highly desirable accomplishment for its workers. The Team of Tests 2 should be given; 60 per cent. must be obtained. If a subject falls below this standard the chances are that she will be more satisfactory in one of the other departments where a low grade of mentality is acceptable.

In the Selecting Department, good judgment and the ability to continue at a monotonous task are best revealed by TeamVI and the Binet Weight Test. Grade C in the memory tests on the Binet Scale in also a requisite for the best grade of worker. Team VII should be given to determine which persons possess special ability for this type of work. The Selectors of Raw Feathers do not need to reach as high a grade of intelligence as do the Selectors of the Finished Feathers and Team III should be given to them. The 3rd or 4th Grade on the Motor Control Graded Scale is sifficiently good. Grade C in the Binet Memory Tests is desirable. There should be no more than seven piles in the Sorting Test, and the time for the performance should not exceed 360″.

In the Boa Department, the Motor Control Test is again the principal criterion by which to judge whether a person can engage to advantage in a monotonous task. Team III or Team VI should be given. Grade C in the Auditory Memory Tests is sufficient.

In the Dyeing Department, Team II is to be given.

The Benders and Twisters should be given Team 2 graded 3rd or 4th in the Motor Test.

The Errand Girls should have B-Grading in the Auditory Momory Tests from the Binet-Scale and be given Team II.

In this schedule it will be noted that the lowest standard in test-rating is acceptable in the Sorting of Raw Feathers, and in the Department of Benders and Twisters. It was found that many girls in these two departments were below average intelligence. None of them had a mental age on the Binet-Scale of more than 10.8., and the majority ranged from 9.4 to 10., while one girl (151) measured but 8.6. The point of importance in this connection is that great many persons who rank below the norm in psychological tests could be usefully employed if a careful study were made not only of the mental equipment of the subjects but also of the degree of mental development required for the work in factories.

Until we know by what mental scale the majority of those engaged in Factory Work should be judged, we cannot

determine any definite criterion for the mental measurements of the steady, unskilled workers of the world. In this study, if we consider the Binet Tests apart from the other standards of measurement, a condition is revealed which permits the conclusion that a great many individuals often regarded as a dead loss to the industrial world can be utilized to good purpose in a number of industries where manual dexterity is more of a requirement than judgment or the ability to reason.

The Experimenter believes that this study has laid the foundation for the utilization of Teams of Tests calculated to measure this grade of labor material.

Furthermore, other questions whose sulution will greatly increase the efficiency of the factory have been solved. The tests have determined whether an applicant will do better work as a selector, a bender, an employee in the Fancy Department, in the Dyeing Department or as an Errand Girl.

When promotions are to be made, the Standards of achievement that are to be considered a legitimate gauge by which to designate the worthiest condidates can be ascertained from an inspection of the analysis of the requirements set forth in the Teams of Tests. Teams II, VI, and VII should be given to those eligible for Promotion.

The fact that the relationship existent between the tests and the Criterion are of a positive nature, and that, in some cases they are of considerable magnitude, are convincing proof of the practical applicability of the methods delineated.

Chapter VIII.

CONFIRMATION OF THE TESTS
Three Years Later.

Calculations in which the human element enters to the degree in which it is present in such tests as those described in this paper are always open to question by the non-scientific who are prone to look upon this sort of analysis as theory pure and simple.

In order to refute any possible charge that the actual progression and evolution of the personnel of the Factory had been in contradiction to the results of the tests and the conclusions drawn therefrom, the Examiner visited the factory after allowing an interval of three years to elapse and carefully ascertained the status of progression or regression of every subject examined three years, earlier who still remained in the employ of the Firm. As much information as possible was gathered regarding those who had left the firm.

It is a source of satisfaction to record that the findings of the tests are vindicated by a fairly close parallel between what the tests predicted and what actually developed in the operation of the factory. The girls scoring the highest mental rating were found to have won promotion to positions of Forewomen, while at the other end of the scale those rating lowest had been dismissed and were no longer employed, or else had remained in statu quo.

Twenty-eight of the original force were found to be still in the employ of the Firm. Of the forty-seven persons who had left, however, thirteen had married, so that they are not to be included in the group which the tests designated as lacking in stability.

It is interesting to note that the mental status revealed by the tests of the subjects leaving to get married is contradictory and inconclusive and permits no such deduction, for instance, as that the more mentally alert have a better matrimonial chance than those ranking lower in the scale. The records show that the subjects achieving matrimony were nearly equally divided between the four quartiles of the scale.

There were, then, thirty-four girls who had left the Factory for other reasons than matrimony.

Of this thirty-four had proved undesirable or undesiring about half had been designated by the tests as unsuited for monotonous work. The Motor Control Test was the principal

index in this diagnosis, although the careless execution of the two Cancellation tests was also noted in this connection.

The distinction between undesirable and undesiring has been made because in the class of unsteadfast persons are some who are in no way mentally lethargic, but are, on the contrary, alert and adventurous, and leave jobs, not through inefficiency but through a craving for larger opportunities. Persons such as No. 48., No. 15., and No. 5. were, it was ascertained, employed as saleswoman, telephone operator and cashier in other concerns. All three of these occupations put a girl on a somewhat higher social basis in the estimation of her comrades, and are also jobs offering an increase in salary.

Subjects No. 2., No. 7., No. 11., No. 27. and No. 42 had left the Factory through other causes than matrimony, or an advance in position. Their performances in the Motor and Cancellation and Sorting Tests shewed them to be incapacitated for monotonous labor and this fact was incorporated in the reports made at the time they were tested.

An Employer taking into consideration the results of the Motor, Cancellation and Sorting Tests would never have given work of a monotonous character or tasks requiring delicate manipulative ability to girls ranking as these girls ranked under tests designated to bring out manual dexterity and the capacity to persevere at a repetitive type of work.

A comparison between the records of the 28 girls who had remained at the Factory and the records of those who had left shows that the first group measured, on an average, 11.11 on the Binet-Simon Scale, while the second group scored an average mental age of 10.85, according to the results of the Binet test. In the Association, Linguistic, Directions and Special Tests, the first group obtained an average score of 73 per cent. with a mean deviation of 6.75., while the second group made an average score of 69.56 per cent., with a mean deviation of 8.9. The differences, though slight, are suggestive and might be of greater significance in a larger group.

The original testing established also what three years of subsequent experience confirmed, namely that subjects No. 65., No. 3., No. 63., No. 24., No. 45 were capable of performing the tasks assigned to them. These girls were found to have progressed in value to the Firm, and in productiveness to themselves.

The most promising subjects, as revealed by high test-scores, of those originally examined, include No. 65., No. 29., No. 24., No. 71., No. 5., No. 75. These subjects never fell below the median in the test-scores in both the Motor and Intelligence tests. They were found to have been advanced to positions of responsibility demanding initiative and alertness of mind. Subject No. 65 is often called upon now to

converse with the Firm's customers and demonstrate the finished product placed on sale.

Subjects No. 24 and 30 who obtained ratings in the first and second quartiles in the tests were occupying positions as Forewomen. Numbers 46., 23., 20., 67., 12., and 42. were continuing the same work in which they had been engaged at the time of their examination, but their wages had been advanced owing to their efficiency and incresed productiveness.

Twelve of the girls remaining in the employ of the Firm upon this re-investigation were found to have been graded generally in the 3rd or 4th Quartiles in their Test-Records.

To those who have made a study of psychological testing, it will not be surprising that such corroboration of the test-findings was obtained both at the top and at the bottom of the Examiner's scale. As is well-known, the tests offer a reliable means of ascertaining the very quick and the very slow.

Not only, however, are the tests of high value in asisting the Employer to arrive at a quick verdict as regards those ranking at the two extremes of the scale but what is even more desirable they provide for him a means of arriving at a fairly accurate estimate of the industrial value of those in the zone between the very good and the very poor.

APPENDIX.

SPECIAL TESTS FOR THIS INDUSTRY.

VI. 1. Judgment Test.

Twenty strips of card-board were cut to approximate the dimensions of the feathers used in the factory, thus:

Feather Dummies.

LENGTH		WIDTH	LENGTH		WIDTH
20.5 in.	x	9.75 in.	12. in.n	x	7.25
17.	x	8.	13.5	x	8.
8.5	x	7.5	15.5	x	8.25
18.	x	7.	14.5	x	17.25
8.5	x	6.125	11.	x	7.
12.5	x	6.5	11.5	x	7.
16.5	x	10.5	14.	x	8.75
18.	x	11.	16.5	x	8.25
18.25	x	9.	10.	x	9.25
15.5	x	8.	22.5	x	11.

Instructions: Each strip of card-board should be cut accurately as per above scale from fairly heavy card-board. An extra strip should be shown to the subject. The subject is told that a number of pieces of card board are to be held before her and that each one has been cut to correspond to the length and width of the feathers which she is in the habit of sorting. The experimenter says further: You are to observe the strip held before you and you are to write your estimate of the length and width of each feather-dummy shown you. There are numbers on the sheet of paper given you that correspond with the numbers on the pieces of card-board.

Each individual was given a sheet of paper at the top of which she was told to write her name and the number that had been given her. Below, under the word "Length", she was to write her approximate estimate of the length of the card, and under the word "Width", the estimated width. A sheet of paper with these headings was given to each girl.

Name.....

Identification Number.....

Length.....Width

1. by

2. by

3.....etc. up to and including 20.

The pieces of card-board were each held at a distance of 18 inches from the subject and exposed for 10″. A stop-watch was used.

Scoring: For each correct judgment, a credit of 5 per cent was given. For every error of a quarter of an inch, 5 per cent was deducted from the score. For every error of more than one-quarter of an inch, one per cent. for each half-inch of error was deducted from the score.

VI. Sorting Test.
Materials: A hundred strips of stiff card-board were cut to correspond with twenty of the most commonly used feathers, similar to those employed in the previous test. Five strips of each size were cut.

Instructions: These strips were laid before the subject in a random heap. The subject was told to treat the pieces of card-board as if they were feathers and sort them into piles, according to size. "Lay those that are of the same length and width together. Work as quickly and as carefully as you can."

Scoring: Score the time required for the entire sorting. Credit 100 per cent. for five distinct piles of cards of like size. For every error add five seconds to the subject's time-score. If more than five piles are made, add five seconds to the total time-score for each additional pile. Subjects are graded on the basis of their final time-score.

VI...3. Card Sorting.

Materials: Two packs of ordinary playing cards from which the two-spots had been removed, as 100 cards are scored more easily than 108 cards.

Instructions: The subject was told to sort the pack of cards into four neat piles according to suits. She was shown a sample card of each suit to preclude any errors on the ground of lack of knowledge of the four suits. If a subject is unfamiliar with the cards note should be made of it. All of these subjects knew the suits and played cards "occasionally."

Each subject was directed to sort the cards as rapidly as possible without making any errors.

Scoring: Time was recorded on the individual record blank under Test VI...3. The majority of errors were corrected by the subject who penalized herself because of the extra time spent in rectifying mistakes. The Munsterberg method of penalties was used. (Ref. 46). The subjects were arranged as in the other tests in rank order of merit.

VI. 4. Discrimination of Color Test.

Materials: From a Wholesale Worsted Firm, a quantity of wools, consisting of defective portions of skeins of all colors and grades, was purchased. There were ten shades each of four different colors, Blue, Green, Purple and Orange. Pieces of these shades were cut into fifteen inch lengths and tied in bows. There were forty other shades and colors which were cut also into fifteen inch lengths and made into bows. Four sets were arranged in each of which were placed a complete set (i. e. ten shades), of either the Blue, Green, Purple or Orange Wools. Mixed with each of these four sets were put seventy-five bits of other worsteds, chosen at random. Each set contained exactly the same number of pieces of worsted. Each of these complete sets was put into a large envelope on the outside of which was written the name of the principal color contained therein. i.e. Blue, Green, Orange or Purple. Before each subject a sheet of 8x10 paper was laid.

Instructions: The directions were given orally, as follows: Each of you has been given a sheet of paper. Write your name on top of it. You have been given an envelope that contains bits of worsted of different colors. Those of you who have an envelope upon which Green is written, are to sort out the bits of worsted that are green; in the same way, those of you who have an envelope bearing the word "Blue" are to sort out all the blue shades of wool, and so on with

the Orange and the Purple. Pay no attention to any color save the one you have been told to select. When the signal "Ready" is given, empty the contents of your envelope on the table and begin at once to sort out your color. Put the worsteds that you have been told to select on the white paper before you. Sort the you are told to Stop, do so at once. Sort the colors as quickly and carefully as you can.

One minute was allowed for the test.

After a girl had sorted one set, she was given another set until each girl had sorted each set.

Scoring: The results of the four color sortings were averaged. There were 85 bits of worsted in each envelope, ten of which were removed from the rest. A credit of 1 was given for each selection of the designated shade, and a deduction of 1 when a color was misplaced. The time was a constant element in all instances, so that the rank order of merit was derived on the basis of accuracy alone.

VI...5. The Motor Control Test.

Materials: On a sheet of 8x11 on legal paper, make lines on the typewriter one-quarter of inch long, approximating thirty-two rows with thirty-eight slants in a row, thus: / / / / /

Instructions: The subject is given a sheet and told at a signal of "Ready" to draw a line through the center of each slant. The Experimenter illustrates. The time allowed for the test is optional with the Experimenter. As this test is intended to determine whether a subject has sufficient stability to persist with steadiness and motor control at an uninteresting and monotonous task, it is desirable to have the sheet completed, or a time-limit of seven minutes and thirty seconds given. The thirty second periods should be checked off by the subject at the command of the Experimenter, who tells the subject to "check" at 30-second intervals. The Experimenter illustrates by showing the subject a check-mark and she is told to continue until directed to "stop".

Scoring: Astandard cross-stroke of approximately an eighth of an inch longer than the obliques was chosen, and only lines which crossed the slant were count-

ed. The number of passable strokes were estimated on this basis. The sheets were graded by three judges who considered them from the standpoint of neatness and precision of stroke, and arranged them first in five piles—Excellent, good, fair, poor ,and bad. The Experimenter then numbered these five piles 1—2—3—4—5. The three judges then arranged the papers in rank order. The average position obtained by each individual in the Judges' Grading was calculated. The scoring was on the basis of the first grading 90 per cent. and over for the Excellent pile; 80 per cent. to 90 per cent. for the good; 65 per cent. to 80 per cent. for the fair; 30 per cent. to 65 per cent. for the "poor" and 30 per cent. and below for the very poor or "bad" papers.

BIBLIOGRAPHY.

1. H. L. Hollingworth.....Vocational Psychology. Pp. 56 to 79. Chapters V., VII., XI.
2. J. Mc.K. Cattell and L. Farrand.....Physical and Mental Measurements of the Students of Columbia University: Ps. R. 3: 1906: pp. 618 to 648.
3. C. Wissler: The Correlation of Mental and Physical Tests: Ps. Mon. 3; No. 6.
4. H. Munsterberg: Psychology and Industrial Efficiency.
5. E. L. Thorndike: The Psychological Methods of Testing Intelligence. P. 210.
   1. Ed. Ps. Vol. III.
   2. An Introduction to the Theory of Mental & Social Measurements.
6. G. M. Whipple: Mental and Physical Tests: Opposites Test: P. 79: Analogies Tests P. 89. Substitution Test: P. 133. Rote Memory P. 160.
7. Vocational Bureau of Boston, Reports
8. Article in London Times, October, 1920.
9. Army Tests: The Personnel System of the Army. Pub. by War Department.
10. Trade Tests: J. Crosby Chapman.
11. W. D. Scott, Advertising and Selling: Oct. 25, 1915.
12. Report of American Tobacco Co., 1916., edited by J. M. Bruce.
13. H. L. Gardner, The Selection Problem of Cheney Bros., U. S. Bur. Labor Statistics, Bul. 227.
14. R. Feiss: Personal Relationship as a basis of Scientific Management. An. Am. Acad. May, 1915.
15. Dallas Consolidated Railway Co. Reports.
16. Curtis Pub. Co.
17. Report of the Bell Telephone Co. 1921.
18. Roy W. Kelly, Hiring the Worker.
19. St. Elmo Lewis, Industrial Management, March 1917.
20. W. F. Kemble, Choosing Employees by Test.
21. Henry C. Link, Employment Psychology.
22. R. A. Spaeth, Industrial Management April 13, 1920.
23. M. Bloomfield, Employment Management, Selected Articles, H. W. Wilson Co. N. Y. City. Employment Problems Ind. Man. August, 1917.
24. F. F. Taylor, Metropolitan Life Ins. Co., Ind. Man. May 1917. P. 873.
25. M. Jaques, Choosing Employees by Test.
26. K. Murdoch, A Sewing Scale.
27. Nat. Ass'n Corp. Schools. Report of Committee on Vocational Guidance, June 1915.
28. Helen T. Wooley, Mental & Physical Meas. of Working Children Mon. 18: 14. June 1915.
29. Bernard Muscio, Lectures on Industrial Psychology.
30. Brewer and Kelly: A Critical Bibliography of Vocational Guidance, 1917. Harvard Press.
31. Vocational Tests, Teachers College.
32. J. Weidensall: The Mentality of the Criminal Woman., Warwick and York. Pp. 142—158. 176.
33. H. Goddard-Binet Measuring Scale of Intelligence: What it is and How it is to be Used. Tr. Sc. 1911.
34. K. B. Davis., Commercialized Prostitution in N. Y. City. Bur. of Social Hygiene.

35. Report of Dr. Olga Bridgman and Dr. L. Morrow, State School for Girls, Geneva, Ill. "Training Sch."; May, 1912.
36. School and Society: June 23, 1917. P. 747.
37. Mental and Sensory Tests: Idaho Industrial School Report.
38. T. H. Haines: The Measurement of Delinquents. Exper. Ps. 1., 1916.
39. E. L. Thorndike: Educational Psy. Vol. 2. P. 260.
40. W. Stern: The Psychological Methods of Testing Intelligence. Ed. Ps. 13., P. 414.
41. A. I. Gates: The Mnemonic Span for Visual and Auditory Digits. J. Exper. Ps. 1919., I. 393-403.
42-43-44. Psychological Mon., No. 13., Pp. 53-55; 75; 85.
44a. Kirkspatrick: Development and Learning: Archives of Ps. 1909., P. 36.
45. Dr. Augusta Bronner: A Comparative Study of the Delinquent Girl.
46. Munsterberg Method of Penalties: Whipple: Mental and Physical Measurement. Page 313.
47. J. Mc.Keen Cattell and L. Farrand: Ps. R. No. 3: 3. 1896., 618-648.
48. W. G. Chambers: Individual Differences in Grammar Grade Children J. Ed. Ps., 1910; Pp. 61-75.
49. J. E. Wallin: Experimental Oral Euthenics: Dental Cosmos: April, May, 1912. Page 323.
50. E. A. Doll: The A-Test with the Feeble-minded: Train Sch. 10; 13. Pp. 49-57.
51. C. Burt: Experimental Tests: Higher Mental Processes and their Relation to General Intelligence. J. E. Pd. 1: (93-112).
52. S. Wyatt: The Quantitative Investigation of Higher Mental Processes. Br. J. Ps. Vol. 6., 1913. Pp. 109-133.
53. E. L. Thorndike: School and Society. Vo; 9., P. 192.
54. Woodworth and Wells: Ps. Mon. 54., 57.
55. E. L. Thorndike: School and Society. Aug. 12, 1916. P. 261.
56. C. Ritter: Ermudungemessungen: Z. Ps. 24: 1900: 401-404.
57. E. L. Thorndike: Educational Psy. Vol. III. P. 370.
58. B. R. Simpson: Correlation of Mental Abilities. Col. Con. Ed. 53; 1912.
59. Stern, Wm. The Psychological Methods of Testing Intelligence Educational Mono. 1914.
60. Burt. Experimental Tests of General Intelligence Brit. Journal Psy. Vol. 3—P. 94.

VITA

The writer was born in Ludlow, N. Y. May 10th, 1880. She was prepared for college at the New York Preparatory School and at Rosemary Hall, and has spent some time in travel and study abroad. She received the degree of A. B,, Barnard College, 1911: A. M. Columbia University, 1912. She specialized in Psychology and Socialogy during her entire college course. The subject of her Master's Thesis was, "An Introspective Study of Association Reaction Times." Since then, has taken courses in Psychology, Psychiatry and Neurology at Cornell University and Medical School, and was assistant in the Department of Neuro-Pathology, 1914 - 1921. She has been a Clinical Psychologist in New York City since 1912: and is now Consulting Psychologist at the Mental Clinic Bellevue Hospital, She is a member of the N. Y. Psychological Association, the N. Y. State Society Consulting Psychologists, and the Psychological Corporation.